

[illegible]

[illegible]

```

REFERENCE Unlabeled.
AUTHORS 1 (bases 1 to 7218)
TITLE Dorner, F., Schellinger, F., and Falkner, F. Gunter.
JOURNAL Recombinant fowlpox virus
FEATURES Patent: US 5670367-A 14-23-SEP-1997;
source Location/Qualifiers
1..7218
/organism="unknown"

BASE COUNT 1944 a 1491 c 1486 g 1929 t 368 others

ORIGIN

Query Match 10.1%; Score 93.2; DB 6; Length 7218;
Best Local Similarity 3.5%; Pred. No. 2,2e-12;
Matches 14; Conservative 257; Mismatches 125; Indels 0; Gaps 0;

364 TTCTCAGTGCCAGCAGACGTTATGAAGCATCTGACCTTAAGCTGGAGAGAGCGGCA 423
1441 TTGTGACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1382
424 GAAGCAAGCGGAGAGAGAGAGAAGAGAAGAGAATGATGAGATGACGATTCAGATTA 483
1381 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1382
484 TCTCTGAGAGCAAAAGCCCTGTCAACAAGTCAAAAGCGTGGTGCOCAGCAAGCGG 1322
1321 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1262
544 AGCTGGCTAGAAAAAAAAGCTGGAAGAAAGAGAGAGAAATAGAGCCGCTTGA 603
1261 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1202
604 GACCAAGCCCTGTAAAAAGGCCAAAGCCACAGCAGCAGCAAGCAAGCCAGATTCAG 663
1201 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 7142
664 AATGAGAGAGCGAGCGCTTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG

```


us-09-844-864-16.rn

RESULT 2
US-08-781-891-209/c
Application US/08781891

GENERAL INFORMATION:
APPLICANT: Fu, Ying-Hui
APPLICANT: Yu, Chang-Fu
APPLICANT: Oshima, Junko
APPLICANT: Mulligan, John T.
APPLICANT: Schellenberg, Gerald D.
TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
TITLE OF INVENTION: WERNER'S SYNDROME
NUMBER OF SEQUENCES: 209
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentL Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,891
FILING DATE: 27-DEC-1996
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: No. 6090620tenburg ph.d., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 240052.419
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
TOW: 209:
INFORMATION FOR SEQ ID NO:
SEQUENCE CHARACTERISTICS:
LENGTH: 51259 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

8

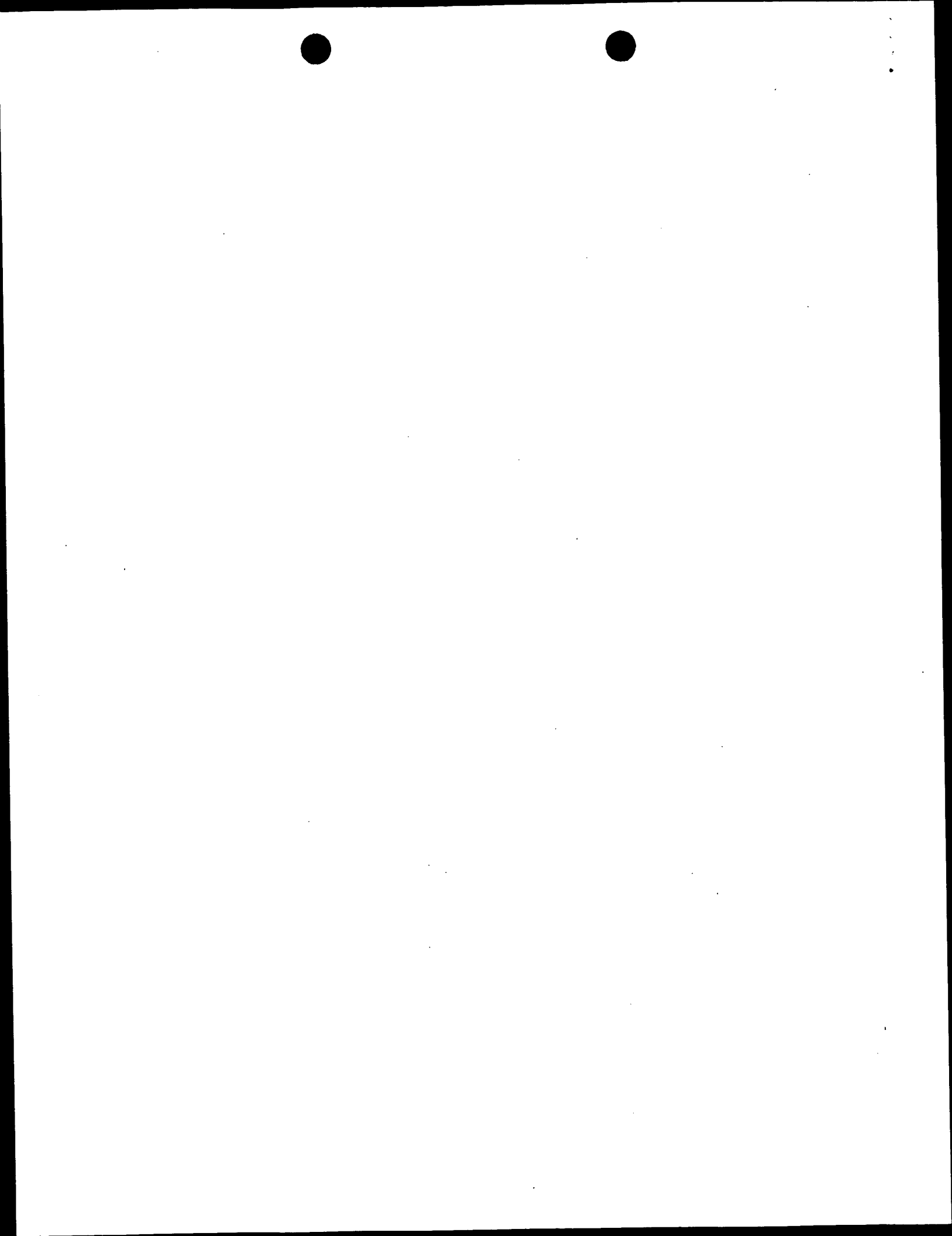
RESULT 3
US-08-781-891-208/c
Sequence 208, Application US/08781891

```

GENERAL INFORMATION:
APPLICANT: Fu, Ying-Hui
APPLICANT: Yu, Chang-En
APPLICANT: Oshima, Junko
APPLICANT: Mulligan, John T.
APPLICANT: Schellenberg, Gerald D.
TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
TITLE OF INVENTION: WERNER'S SYNDROME
NUMBER OF SEQUENCES: 209
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED and BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,891
FILING DATE: 27-DEC-1996
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: No. 6090620tenburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 240052.419
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 208:
SEQUENCE CHARACTERISTICS:
LENGTH: 16442 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-781-891-208

Query Match 6.4%; Score 59.4; DB 3; Length 16442;
Best Match Similarity 57.6%; Pred.No.1.1e-05;
Matches 125; Conservative 0; Mismatches 91; Indels 1; Gaps

```



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 24, 2003, 05:06:35 : Search time 77 seconds

(without alignments)
3680.121 Million cell updates/sec

Title: US-09-844-864-16

Perfect score: 924
Sequence: 1 cagccgcctctctcgcgcg.....tttcgagcgaagcttatg 924

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database:

- 1: Issued_Patents_NA.*
- 2: /cgn2_6/ptodata/2/1na/5A_COMB.seq.*
- 3: /cgn2_6/ptodata/2/1na/5B_COMB.seq.*
- 4: /cgn2_6/ptodata/2/1na/6A_COMB.seq.*
- 5: /cgn2_6/ptodata/2/1na/6B_COMB.seq.*
- 6: /cgn2_6/ptodata/2/1na/6C_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	93.2	10.1	7218	1 US-08-232-463-14	Sequence 14, Appl
C 2	63.4	6.9	51259	3 US-08-781-881-209	Sequence 209, App
C 3	59.4	6.4	16442	2 US-08-781-891-208	Sequence 208, App
C 4	59.2	6.4	3489	2 US-08-728-323A-1	Sequence 1, Appl
C 5	59.2	6.4	3489	4 US-09-298-568-1	Sequence 1, Appl
C 6	59.2	6.4	32207	2 US-08-770-379-20	Sequence 20, Appl
C 7	59.2	6.4	32207	4 US-08-757-669A-20	Sequence 20, Appl
C 8	59.2	6.4	32207	4 US-09-230-371A-20	Sequence 20, Appl
C 9	57.6	6.2	3211	2 US-08-574-959A-8	Sequence 8, Appl
C 10	57.6	6.2	3211	4 US-09-357-014-8	Sequence 8, Appl
C 11	57.6	6.2	3901	2 US-08-574-959A-6	Sequence 6, Appl
C 12	57.6	6.2	3901	4 US-09-357-014-6	Sequence 6, Appl
C 13	56.6	6.1	2518	4 US-09-433-699-3	Sequence 3, Appl
C 14	55.2	6.0	2340	3 US-09-022-983-4	Sequence 4, Appl
C 15	55.2	6.0	2477	4 US-09-490-692-3	Sequence 3, Appl
C 16	54.4	5.9	9636	4 US-08-323-170B-1	Sequence 1, Appl
C 17	54.4	5.9	9636	4 US-08-954-441-1	Sequence 1, Appl
C 18	54.4	5.8	1236	4 US-08-741-134-5	Sequence 5, Appl
C 19	53.2	5.6	289	4 US-08-007-005-17	Sequence 17, Appl
C 20	52.2	5.6	289	4 US-09-244-786-17	Sequence 17, Appl
C 21	52.2	5.6	15378	3 US-08-785-420-1	Sequence 1, Appl
C 22	51.8	5.6	966	2 US-08-766-738-2	Sequence 2, Appl
C 23	51.8	5.6	966	4 US-09-262-610-2	Sequence 2, Appl
C 24	51.6	5.5	489	1 US-07-879-685B-3	Sequence 3, Appl
C 25	50.6	5.5	5183	1 US-08-459-568-3	Sequence 3, Appl
C 26	50.6	5.5	5183	1 US-08-399-411-3	Sequence 3, Appl
C 27	50.6	5.5	5868	3 US-08-516-859A-3	Sequence 3, Appl

28	50.6	5.5	5868	4 US-09-586-472-3	Sequence 3, Appl
29	50.6	5.5	5868	4 US-09-528-706-3	Sequence 3, Appl
30	50.4	5.5	1926	4 US-09-249-585A-4	Sequence 4, Appl
C 31	50.4	5.5	1931	2 US-09-130-114-2	Sequence 2, Appl
C 32	50	5.4	258	4 US-09-345-882-21	Sequence 21, Appl
C 33	50	5.4	6002	4 US-09-345-882-4	Sequence 4, Appl
C 34	50	5.4	162450	4 US-09-345-882-1	Sequence 1, Appl
C 35	49.4	5.3	3100	1 US-08-296-362-1	Sequence 1, Appl
C 36	49	5.3	2188	1 US-07-865-662E-10	Sequence 10, Appl
C 37	49	5.3	2188	4 US-08-374-219B-10	Sequence 10, Appl
C 38	48.8	5.3	1678	3 US-08-650-766-2	Sequence 2, Appl
C 39	48.8	5.3	1954	3 US-08-922-635-2	Sequence 2, Appl
C 40	48.8	5.3	2255	2 US-08-741-134-1	Sequence 1, Appl
C 41	48.8	5.3	3318	3 US-08-650-766-3	Sequence 3, Appl
C 42	48.8	5.3	3318	3 US-08-922-635-3	Sequence 3, Appl
C 43	48.8	5.3	3385	3 US-08-650-766-1	Sequence 1, Appl
C 44	48.8	5.3	3385	3 US-08-922-635-1	Sequence 1, Appl
C 45	48.8	5.3	15202	3 US-08-922-635-21	Sequence 21, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14/C
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ. ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZ9pt-F1s
US-08-232-463-14
Query Match 10.1%; Score 93.2; DB 1; Length 7218;

Best Local Similarity 3.5%; Pred. No. 1.5e-14;
Matches 14; Conservative 257; Mismatches 125; Indels 0; Gaps 0;

```

QY 364 TTCTCAGTGGCCAGCAAGCTTATGACATCAGACTTAACCTGGAGGAGAGAGAGAA 423
    || | : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1441 TTGCTACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1382
QY 424 GAAGAAAGGAGAGAGAGAGAGAGAGAGAGATGATGAGGATGAGATGAGATGATA 483
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1381 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1322
QY 484 TCTCGAGAGAGCAAGCCCTGTCAACAGTCAAGAGCTGCTGCTGCTGCTGCTGCTG 543
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1321 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1262
QY 544 AGCGTGGTCAAGAAAAAGCTGGAAGAAAGAGAGAAATAGACCAAGCTGTGA 603
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1261 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1202
QY 604 GACAGAGCCCTGTGAAAAAGCCCAAGCCAGCCAGCAAGCAAGCAAGCAAGTTCAG 663
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1201 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1142
QY 664 AATGAGAGAGCCAGCTTGGGGGAGAGTGAAGTGGCTTCCCTGGCTGTGCTG 723
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1141 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1082
QY 724 CAGGACAGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 759
    : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1081 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1046

```

RESULT 2

```

US-08-781-891-209/c
; Sequence 209, Application US/08781891
; Patent No. 6090620
; GENERAL INFORMATION:
; APPLICANT: Fu, Ying-Hui
; APPLICANT: Oshima, Junko
; APPLICANT: Mulligan, John T.
; APPLICANT: Schellenberg, Gerald D.
; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
; TITLE OF INVENTION: WERNER'S SYNDROME
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,891
; FILING DATE: 27-DEC-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6090620tenburg Ph.D., Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 240052.419
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 209:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 51259 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

```

TOPOLOGY: linear
US-08-781-891-209

Query Match 6.9%; Score 63.4; DB 3; Length 51259;
Best Local Similarity 59.9%; Pred. No. 1.6e-06;
Matches 106; Conservative 0; Mismatches 71; Indels 0; Gaps 0;

```

QY 408 GGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATGATGAGA 467
    ||||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 257 GGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATGATGAGA 198
QY 468 TGAAGATGACATATATCTGTGAGAGCAAGCCCTGTCAACAGTCAAAAGCTGCT 527
    ||||| | | | | | | | | | | | | | | | | | | | | | | | |
Db 197 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATGATGAGA 138
QY 528 GCCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 584
    | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 137 GAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 81

```

RESULT 3

```

US-08-781-891-208/c
; Sequence 208, Application US/08781891
; Patent No. 6090620
; GENERAL INFORMATION:
; APPLICANT: Fu, Ying-Hui
; APPLICANT: Oshima, Junko
; APPLICANT: Mulligan, John T.
; APPLICANT: Schellenberg, Gerald D.
; TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
; TITLE OF INVENTION: WERNER'S SYNDROME
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/781,891
; FILING DATE: 27-DEC-1996
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: No. 6090620tenburg Ph.D., Carol
; REGISTRATION NUMBER: 39,317
; REFERENCE/DOCKET NUMBER: 240052.419
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 208:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 16442 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-781-891-208

```

Query Match 6.4%; Score 59.4; DB 3; Length 16442;
Best Local Similarity 57.6%; Pred. No. 1.1e-05;
Matches 125; Conservative 0; Mismatches 91; Indels 1; Gaps 1;

```

QY 376 CAGGACGTTATGACATCAGACCTTAACCTGGAGGAGAGAGAGAGAGAGAGAGAG 435
    ||||| | | | | | | | | | | | | | | | | | | | | | | | |
Db 16369 CAGGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 16310
QY 436 GAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATGATGATGATG 495

```


? PATENT NO.: 5849564
 ? GENERAL INFORMATION:
 ? APPLICANT: Chang, Yuan
 ? APPLICANT: Bohenzky, Roy A.
 ? APPLICANT: Russo, James J.
 ? APPLICANT: Edelman, Isidore S.
 ? APPLICANT: Moore, Patrick S.
 ? TITLE OF INVENTION: POLYPEPTIDES FROM KAPOSI'S SARCOMA-ASSOCIATED
 ? HEPHERSVIRUS, DNA ENCODING SAME AND USES THEREOF

```

1 STREET 1185 Avenue of the Americas
2 CITY: New York
3 STATE: New York
4 COUNTRY: U.S.A.
5 ZIP: 10036
6
7 COMPUTER READABLE FORM:
8 MEDIUM TYPE: Floppy disk
9 COMPUTER: IBM PC compatible
10 OPERATING SYSTEM: PC-DOS/MS-DOS
11 SOFTWARE: PatentIn Release #1.0, Version #1.30
12 CURRENT APPLICATION DATA:
13 APPLICATION NUMBER: US/08/757,669A
14 FILING DATE:
15 CLASSIFICATION: 424
16
17 ATTORNEY/AGENT INFORMATION:
18 NAME: White, John P.
19 REGISTRATION NUMBER: 28,678
20 REFERENCE/DOCKET NUMBER: 45185-F
21 TELECOMMUNICATION INFORMATION:
22 TELEPHONE: (212) 278-0400
23 TELEFAX: (212) 391-0525
24
25 INFORMATION FOR SEQ ID NO: 20:
26 SEQUENCE CHARACTERISTICS:
27 LENGTH: 32207 base pairs
28 TYPE: nucleic acid
29 STRANDEDNESS: double
30 TOPOLOGY: linear
31
32 MOLECULE TYPE: DNA (genomic)
33
34 US-08-757-669A-20

```

```

Query Match 6.4%; Score 59.2; DB 4; Length 32207;
Best Local Similarity 52.8%; Pred. No. 1.6e-05;
Matches 151; Conservative 0; Mismatches 133; Indels 2; Gaps 1

QY 408 GGAGGAGGAGAGAGAAAGAAAGGAGGAGGAGAGAGAGAGAGAGAGAGAGATGATGACGA 467
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 20836 GGATGACGAGGAGAGATGAGACGAGGAGATGACGAGGAGATGACGAGGAGATGACGAGGA 20837

QY 468 TGAAGATGACGATATATCTCTGGAGGAGCAAAAGCCCTGTCAAACAATGCAAAAGCGTGGT 527
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 20836 GGAGGACGAGGAGAGAGAGACGACGAGGAGGAGGACGAGAGAGAGAGGACGAGGATGACGA 20777

QY 528 GCCCCAGAGAGAGGCGAGCCCTGCTGAAGAAAAAAAAGCTGAAAAAAGAGAGAGAAAT 587
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 20776 TGATGAGAGACATATGAGACGACGAGAGATATACGAGGAGGAGACAAAGAGGAGGACGAGGA 20717

QY 588 AAGAGCCAGCGGTTAGAGACAAAGACCCCTGTGA--AAAGGCCAAAGCCACAGCCAGAGCC 645
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 20716 GGAGGAGGCGCATGTAAGAAACAAAGCTTTGAGCATCCAAATTCACAAACGACGACGAGGCC 20657

QY 646 AAGAGCCAGGATTCAGAAATGAGGAGGCGACGCGCTTGGGGGGGCGAC 691
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 20656 ACAACGACGAGGAGCCACAGCAGCAGGAGGACACAGCAGCAGGAGGCC 20611

RESULT 8
US-09-230-371A-20/c
; Sequence 20, Application US/09230371A
; Patent No. 6348586
; GENERAL INFORMATION:
; APPLICANT: Chang, Yuan
; APPLICANT: Bohenzky, Roy A
; APPLICANT: Russo, James J
; APPLICANT: Edelman, Isidore S
; APPLICANT: Moore, Patrick S
; TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SAROMA VIRUS SEQUENCES AND
; FILE OF INVENTION: US5-PCT-US
; FILE REFERENCE: 45185-G-PCT-US
; CURRENT APPLICATION NUMBER: US/09/230,371A
; CURRENT FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: PCT/US97/13346
; PRIOR FILING DATE: 1997-07-22
; NUMBER OF SEQ ID NOS: 30

```


RESULT 14
 US-09-022-983-4
 : Sequence 4, Application US/09022983
 : Patent No. 6159731
 :
 : GENERAL INFORMATION:
 :
 : APPLICANT: Yang, Xiaolu
 : APPLICANT: Khosravi-Far, Roya
 : APPLICANT: Chang, Howard Y.
 : APPLICANT: Baltimore, David
 : TITLE OF INVENTION: DAXX, A NOVEL FAS-BINDING
 : TITLE OF INVENTION: PROTEIN THAT ACTIVATES JNK AND APOPTOSIS
 : NUMBER OF SEQUENCES: 5
 :
 : CORRESPONDENCE ADDRESS:
 : ADDRESSEE: Wolf, Greenfield & Sacks, P.C.
 : STREET: 600 Atlantic Avenue
 : CITY: Boston
 : STATE: MA
 : COUNTRY: USA
 : ZIP: 02210
 : COMPUTER READABLE FORM:
 : MEDIUM TYPE: Diskette
 : COMPUTER: IBM Compatible
 : OPERATING SYSTEM: DOS

```

RESULT 15
US-09-490-692-3
: Sequence 3, Application US/09490692
: Patent No. 6180353
:
: GENERAL INFORMATION:
: APPLICANT: Nicholas M. Dean
: APPLICANT: Lex M. Cowsett
: TITLE OF INVENTION: ANTISENSE MODULATION OF DAXX EXPRESSION
: FILE REFERENCE: RTS-0120
: CURRENT APPLICATION NUMBER: US/09/490,692
: CURRENT FILING DATE: 2000-01-24
: NUMBER OF SEQ. ID NOS. 176
: SEQ. ID NO 3
:
: LENGTH: 2477
:
: TYPE: DNA
:
: ORGANISM: Homo sapiens
:
: FEATURE:
: NAME/KEY: CDS
: LOCATION: (116)...(2338)
: US-09-490-692-3

```


GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 23, 2003, 14:06:24 ; Search time 45576 Seconds

(without alignments)
590.025 Million cell updates/sec

Title: US-09-844-864-16

Perfect score: 924

Sequence: 1 cagccgcgtctctgcgcgcg.....tttcgagccgaagctatg 924

Scoring table:

IDENTITY-NUC
Gapop 10.0, Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

```

GenEmbl:
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_com:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_scs:*
28: em_un:*
29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pla:*
35: em_hcg_rod:*
36: em_hcg_mam:*
37: em_hcg_vtc:*
38: em_sy:*
39: em_hcgo_hum:*
40: em_hcgo_mus:*
41: em_hcgo_other:*

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	883	95.6	1874	9 AK094267	AK094267 Homo sapl
2	337	36.5	423	6 AX321879	AX321879 Sequence
3	262.4	28.4	156347	2 AC091171	AC091171 Homo sapl
4	262.4	28.4	166878	2 AC023288	AC023288 Homo sapl
5	262.4	28.4	181714	2 AC090442	AC090442 Homo sapl
6	262.4	28.4	184050	2 AC087819	AC087819 Homo sapl
7	262.4	28.4	184444	2 AC090471	AC090471 Homo sapl
8	171.6	18.6	86574	9 HS83387	AL008637 Human DNA
9	122	13.2	135206	2 AC108486	AC108486 Homo sapl
10	120.4	13.0	149940	9 AC093557	AC093557 Homo sapl
11	99.6	10.8	646	5 XLN0PLR	Y00204 Xenopus lae
12	96.2	10.4	800	5 XLN0PLR2	X04766 Xenopus lae
13	93.2	10.1	7218	6 I66494	I66494 Sequence 14
14	90.6	9.8	198421	2 AC125180	AC125180 Mus muscu
15	83.2	9.0	190604	2 AC122303	AC122303 Mus muscu
16	81.6	8.8	200052	2 AL772398	AL772398 Mus muscu
17	81.4	8.8	251076	2 AC097410	AC097410 Rattus no
18	81.2	8.8	2081	9 AK056804	AK056804 Homo sapl
19	79.8	8.6	43553	9 AL357372	AL357372 Human DNA
20	79	8.5	222259	2 AC113078	AC113078 Mus muscu
21	78.4	8.5	241666	2 AC118543	AC118543 Mus muscu
22	77.4	8.4	182740	2 AC115811	AC115811 Mus muscu
23	76.8	8.3	165077	10 AC084382	AC084382 Mus muscu
24	76.2	8.3	170032	2 AC116673	AC116673 Mus muscu
25	76.2	8.2	70391	2 AC110410	AC110410 Rattus no
26	75.8	8.2	144328	9 AC009695	AC009695 Homo sapl
27	75.8	8.2	163495	9 AC022716	AC022716 Homo sapl
28	75.8	8.2	170631	9 AC020751	AC020751 Homo sapl
29	75.8	8.2	170807	9 AC023812	AC023812 Homo sapl
30	75.8	8.2	179803	9 AC068314	AC068314 Homo sapl
31	75.4	8.2	168856	2 AF466883	AF466883 Mus muscu
32	75.4	8.2	169153	2 AC128501	AC128501 Rattus no
33	75.2	8.1	110000	2 AC098456_2	Continuation (3 of
34	75.2	8.1	303943	2 AC127311	AC127311 Mus muscu
35	74.8	8.1	177716	9 AC026188	AC026188 Homo sapl
36	74.6	8.1	152763	2 AC119715	AC119715 Rattus no
37	74.6	8.1	173456	2 AC115183	AC115183 Rattus no
38	74.4	8.1	147670	10 AC084020	AC084020 Mus muscu
39	74.4	8.1	205606	10 AL596204	AL596204 Mouse DNA
40	73.8	8.0	184370	2 AL589845	AL589845 Mus muscu
41	73.8	8.0	120044	2 AC129042	AC129042 Rattus no
42	73.8	8.0	208844	2 AC125070	AC125070 Mus muscu
43	73.6	8.0	215938	2 AC102646	AC102646 Mus muscu
44	73.6	8.0	93682	10 AL591805	AL591805 Mouse DNA
45	73.6	8.0	154733	2 AC131221	AC131221 Rattus no

ALIGNMENTS

RESULT 1									
AK094267									
LOCUS	AK094267	1874 bp	mRNA	linear				PRI 15-JUL-2002	
DEFINITION	Homo sapiens cDNA FLJ36948 fls, clone BRACE2005719, <u>weakly similar</u> to NUCLEOPLASMIN.								
ACCESSION	AK094267								
VERSION	AK094267.1	GI:21753294							
KEYWORDS	oligo capping; fls (full insert sequence).								
SOURCE	Homo sapiens cerebellum cDNA to mRNA, clone_1lb:BRACE2 clone:BRACE2005719.								
ORGANISM	Homo sapiens								
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.								
AUTHORS	1 Tanigami,A., Fujiwara,T., Shibahara,T., Goto,Y., Hiroo,M.,								

Shimizu, F., Wakebe, H., Ono, T., Hishigaki, H., Watanabe, T., Ozaki, K., Sugiyama, T., Irie, R., Otsuki, F., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Nagatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahari, K., Masuno, Y., Nagai, K. and Isogai, T.
NEBO human cDNA sequencing project

Unpublished
2 (bases 1 to 1874)
Isogai, T. and Yamamoto, J.
Direct Submission
Submitted (04-JUL-2002) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0812, Japan
(E-mail: genomics@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
NEBO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'-83'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

FEATURES
location/Qualifiers
1..1874
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="BRACE2005719"
/tissue_type="cerebellum"
/clone_lib="BRACE2"
/note="cloning vector: PME18SFL3"

BASE COUNT 394 a 570 c 604 g 306 t

ORIGIN

Query Match 95.6%; Score 883; DB 9; Length 1874;
Best Local Similarity 100.0%; Pred. No. 3.2e-208;
Matches 883; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGCCCGCTTCTCTCCCGGAGCCATGATCTCATCTACGCCAGTACGACGAGGAAAG 60
DB 992 CAGCCCGCTTCTCTCCCGGAGCCATGATCTCATCTACGCCAGTACGACGAGGAAAG 1051

QY 61 GGAGTGACGACCGCTCTGGGGCTGCGACGTCAGTACGAGAGGCGGACTTGGACCTTC 120
DB 1052 GGAGTGACGACCGCTCTGGGGCTGCGACGTCAGTACGAGAGGCGGACTTGGACCTTC 1111

QY 121 AGACCCGACCTGAGGAGGAGAGAGAGCTGAGGCTGTTCTTCAATGATTTGCTTGGGG 180
DB 1112 AGACCCGACCTGAGGAGGAGAGAGAGCTGAGGCTGTTCTTCAATGATTTGCTTGGGG 1171

QY 181 GAGAAAGCCAAAGAGAGATGATCGCGTGGAGATCCGCCCGCAAGACAGAGAGAG 240
DB 1172 GAGAAAGCCAAAGAGAGATGATCGCGTGGAGATCCGCCCGCAAGACAGAGAGAG 1231

QY 241 AAGAAGATGACGCGGTCACATTCCTCCATCCAGGCTCAGTCTCCCATGCTTC 300
DB 1232 AAGAAGATGACGCGGTCACATTCCTCCATCCAGGCTCAGTCTCCCATGCTTC 1291

QY 301 ATGTGAGAGTCAAGTCTTCTCCCAATTTACTTTTCAAGCTCCGGGCTGAGGACCC 360
DB 1292 ATGTGAGAGTCAAGTCTTCTCCCAATTTACTTTTCAAGCTCCGGGCTGAGGACCC 1351

QY 361 GTGTTCTTCAAGTGGCAGAGAGAGTATGAAGATCAGACTTAAGTGGAGAGAGAGAG 420
DB 1352 GTGTTCTTCAAGTGGCAGAGAGAGTATGAAGATCAGACTTAAGTGGAGAGAGAGAG 1411

QY 421 GAGAAGAGGAGGAGAGAGAGAGAGAGAGAGAGATGATGAGATGAGATGAGATGAGAT 480
DB 1412 GAGAAGAGGAGGAGAGAGAGAGAGAGAGAGAGAGATGATGAGATGAGATGAGATGAGAT 1471

QY 481 ATATCTCTGAGAGAGCAAGCCCTGTCAACAAGTCAAAAGGCTGCTGCCAGAGAGAG 540
DB 1472 ATATCTCTGAGAGAGCAAGCCCTGTCAACAAGTCAAAAGGCTGCTGCCAGAGAGAG 1531

QY 541 GCGAGCGTGGCTAAGAAAAAGCTGAAAAAGAGAGAGAAATAGAGCCAGCGTT 600
DB 1532 GCGAGCGTGGCTAAGAAAAAGCTGAAAAAGAGAGAGAAATAGAGCCAGCGTT 1591

QY 601 AGAGACAGAGAGCCCTGTGAAAAAGGCCAACACAGCCAGAGCCAGAGATTC 660
DB 1592 AGAGACAGAGAGCCCTGTGAAAAAGGCCAACACAGCCAGAGCCAGAGATTC 1651

QY 661 AAGAAATGAGAGAGCCAGCTTGGGGGAGACGGTGCAGAGTGGGCTTCCCTGGGCTGTG 720
DB 1652 AAGAAATGAGAGAGCCAGCTTGGGGGAGACGGTGCAGAGTGGGCTTCCCTGGGCTGTG 1711

QY 721 CTGACGACAGAGGCTGCTCCCTGTCACGCCCTTCACCTTGTCTGTAATCAACAGAGGAGT 780
DB 1712 CTGACGACAGAGGCTGCTCCCTGTCACGCCCTTCACCTTGTCTGTAATCAACAGAGGAGT 1771

QY 781 TTGCGGGGGGACATGAGAGCCCTCAGCCCAACTCTCCTCACTTTCAGAGAGCCCACT 840
DB 1772 TTGCGGGGGGACATGAGAGCCCTCAGCCCAACTCTCCTCACTTTCAGAGAGCCCACT 1831

QY 841 GAAGAGCCCACTCTGGGGTCAACAATTAAGTGGCTGTCAGG 883
DB 1832 GAAGAGCCCACTCTGGGGTCAACAATTAAGTGGCTGTCAGG 1874

RESULT 2
AX321879 423 bp DNA linear PAT 15-DEC-2001
LOCUS
DEFINITION
Sequence 410 from Patent WO0172295.
AX321879
VERSION
AX321879.1 GI:17906455
KEYWORDS
human.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
1
Reed, S.G., Lodes, M.J., Mohamath, R., Secrist, H., Benson, D.R., Indrias, C.Y., Henderson, R.A., Flinn, S.P., Algate, P.A., Elliot, M., Mannion, J. and Kalos, M.D.
Compositions and methods for the therapy and diagnosis of lung cancer
Patent: WO 0172295-A 410 04-OCT-2001;
CORIXA CORPORATION (US)
location/Qualifiers
1..423
/organism="Homo sapiens"
/db_xref="taxon:9606"

FEATURES
source
BASE COUNT 84 a 138 c 125 g 76 t

ORIGIN

Query Match 36.5%; Score 337; DB 6; Length 423;
Best Local Similarity 100.0%; Pred. No. 6.5e-73;
Matches 337; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CAGCCCGCTTCTCTGCCCGAGCCATGATCTAGTACGCCAGTACGAGGAGAAAG 60
DB 87 CAGCCCGCTTCTCTGCCCGAGCCATGATCTAGTACGCCAGTACGAGGAGAAAG 146

QY 61 GCACTGACGACCGTCTCTGGGGCTGCGAGCTCAGTACGAGAGGCGGACTTGGACCTTC 120
DB 147 GCACTGACGACCGTCTCTGGGGCTGCGAGCTCAGTACGAGAGGCGGACTTGGACCTTC 206

QY 121 AGACCCGACCTGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
DB 207 AGACCCGACCTGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 266

QY 181 GAGAAGCCAAAGAGAGAGATGATCGGTGAGATCTTCCGCCAGCAACCAAGAGAGAG 240
DB 267 GAGAAGCCAAAGAGAGAGATGATCGGTGAGATCTTCCGCCAGCAACCAAGAGAGAG 326

QY 241 AAGAAGATGACGCGGTTCACCATTTGCTCACTCAGGCTCAAGTCTCCCATGATGCTCC 300

KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
 SOURCE Homo sapiens.
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 REFERENCE 1 (bases 1 to 166878)
 AUTHORS Abola, A.P., Bruno, D., Conn, L., Dela Rosa, M., Faulkner, D., Federspiel, N., Glukhov, S., Hansen, N., Hyman, R., Mao, J., Marathe, R., Morehouse, A.J., Oefner, P., Palm, C.J., Ramirez, D., Southwick, A.M., Wilhelm, J., Yu, S. and Davis, R.W.
 JOURNAL Unpublished
 REFERENCES 2 (bases 1 to 166878)
 AUTHORS Bruno, D., Conn, L., Dela Rosa, M., Faulkner, D., Federspiel, N., Glukhov, S., Hansen, N., Hyman, R., Mao, J., Marathe, R., Morehouse, A.J., Oefner, P., Palm, C.J., Ramirez, D., Southwick, A.M., Wilhelm, J., Yu, S. and Davis, R.W.
 TITLE Submitted (11-FEB-2000) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA
 COMMENT On Jul 19, 2000 this sequence version replaced 91:8980905.
 ----- Genome Center
 Center: Stanford DNA Sequencing and Technology Development Center
 Center code: SDSSTD
 Web site: <http://sequence-www.stanford.edu/group/human/>
 Contact: hum-info@sequence.stanford.edu
 ----- Project Information
 Center project name: 836
 Center clone name: RP11-67H12
 ----- Summary Statistics
 Sequencing Vector: M13mp18; X02513
 Chemistry: Dye-primer; 0% of reads
 Assembly program: Phrap; version 0.990319
 Consensus quality: 151238 bases at least Q40
 Consensus quality: 155909 bases at least Q30
 Consensus quality: 157935 bases at least Q20
 Insert size: 116494; agarose-fp
 Insert size: 165478; sum-of-contigs
 Quality coverage: 9.4x in Q20 bases; agarose-fp
 Quality coverage: 6.6x in Q20 bases; sum-of-contigs
 NOTE: This is a 'working draft' sequence. It currently consists of 15 contigs. The true order of the pieces is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.
 1 2175: contig of 2175 bp in length
 2176 2275: gap of unknown length
 2276 4472: contig of 2197 bp in length
 4473 4572: gap of unknown length
 4573 6489: contig of 1917 bp in length
 6490 6590: gap of unknown length
 6590 8363: contig of 1774 bp in length
 8364 8463: gap of unknown length
 8464 10644: contig of 2181 bp in length
 10645 10744: gap of unknown length
 10745 13373: contig of 2629 bp in length
 13374 13473: gap of unknown length
 13474 17686: contig of 4213 bp in length
 17687 17787: gap of unknown length
 17788 26030: contig of 8244 bp in length
 26031 26131: gap of unknown length
 26132 33532: contig of 7402 bp in length
 33533 33633: gap of unknown length
 33634 46267: contig of 12635 bp in length
 46268 46368: gap of unknown length
 46369 61796: contig of 15429 bp in length
 61797 61896: gap of unknown length

FEATURES
 source
 1. 166878
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="8"
 /clone="RP11-67H12"
 /clone_id="RP11-67H12"
 /clone_lib="RP11 human BAC library 11"
 1. 2175
 /note="assembly_name:Contig42"
 2276. 4472
 /note="assembly_name:Contig24"
 4573. 6489
 /note="assembly_name:Contig25"
 6590. 8363
 /note="assembly_name:Contig28"
 8464. 10644
 /note="assembly_name:Contig31"
 10745. 13373
 /note="assembly_name:Contig33"
 13474. 17686
 /note="assembly_name:Contig34"
 17787. 26030
 /note="assembly_name:Contig35"
 26131. 33532
 /note="assembly_name:Contig36
 clone_end:SP6
 vector_side:left"
 33633. 46267
 /note="assembly_name:Contig37"
 46368. 61796
 /note="assembly_name:Contig38"
 61897. 78516
 /note="assembly_name:Contig39"
 78617. 97532
 /note="assembly_name:Contig40
 clone_end:T7
 vector_side:right"
 97633. 132687
 /note="assembly_name:Contig41"
 132788. 166878
 /note="assembly_name:Contig42"
 BASE COUNT 41842 a 40044 c 40620 g 42959 t 1413 others
 ORIGIN
 Query Match 28.4%; Score 262.4; DB 2: length 166878;
 Best Local Similarity 99.6%; Pred. No. 3.2e-54;
 Matches 263; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 621 AAAGGCCAAGCCACAGCCAGCCAGCAAGCAAGCAATTCAGAAATGAGAGCCAGCC 680
 Db 146754 ACAGGCCAAGCCACAGCCAGCCAGCAAGCAAGCAATTCAGAAATGAGAGCCAGCC 146813
 QY 681 TTGGGGGGGACGGTGCAGAAATGGGCTCCCTGGGCTGCTCCAGGACAGGGTCCGCC 740
 Db 146814 TTGGGGGGGACGGTGCAGAAATGGGCTCCCTGGGCTGCTCCAGGACAGGGTCCGCC 146873
 QY 741 TGTCCAGCCCTCCACCTGTGTGTAATGCAGAGGGGTGTCGGGGCCAAACATGAGAG 800
 Db 146874 TGTCCAGCCCTCCACCTGTGTGTAATGCAGAGGGGTGTCGGGGCCAAACATGAGAG 146933
 QY 801 CCCCTACCCCAACTCTCCACTTTCAGAGAGCCCAAGTGAAGAGCCCACTCGGGGT 860
 Db 146934 CCCCTACCCCAACTCTCCACTTTCAGAGAGCCCAAGTGAAGAGCCCACTCGGGGT 146993
 QY 861 CACATTAAGTTGCTGTGTCAGGA 884
 |||||||||||||||||||||||||||

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS

1 (bases 1 to 184050)
 Birren,B., Linton,L., Nusbaum,C. and Lander,E.
 Homo sapiens chromosome 8, clone RP11-507M15
 Unpublished
 2 (bases 1 to 184050)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,
 Barrn,N., Bastien,V., Boguslavsky,L., Boukhgalter,B., Brown,A.,
 Camarata,J., Campopiano,A., Choquet,Y., Colangelo,M., Collins,S.,
 Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,
 Dodge,S., Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J.,
 Gardna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,
 Hages,B., Hearford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
 Jones,C., Karatas,A., Labrecque,K., Lamazares,R., Landers,T.,
 Lhoczek,J., Levine,R., Liu,G., Maclean,C., Macdonald,P.,
 Marquis,N., Mathews,C., McCarthy,N., McEwan,P., McKernan,K.,
 McPheters,R., Meldrim,J., Meneus,L., Minova,T., Mlenga,V.,
 Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,
 O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
 Phunhkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,
 Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,
 Roy,A., Santos,R., Schauer,S., Schupbach,R., Seaman,S., Severy,P.,
 Sougnuez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,
 Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,
 Travers,M., Travis,N., Trigliio,J., Vassiliev,H., Viel,R., Vo,A.,
 Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,
 Zembek,L., Zimmer,A. and Zody,M.

FEATURES

source

* 32849 34030: contig of 1182 bp in length
 * 34031 34130: gap of 100 bp
 * 34131 35157: contig of 1027 bp in length
 * 35158 35257: gap of 100 bp
 * 35258 72407: contig of 37150 bp in length
 * 72408 72507: gap of 100 bp
 * 72508 75697: contig of 3190 bp in length
 * 75698 75797: gap of 100 bp
 * 75798 78108: contig of 2211 bp in length
 * 78109 78208: gap of 100 bp
 * 78209 82150: contig of 3942 bp in length
 * 82151 82250: gap of 100 bp
 * 82251 88573: contig of 6323 bp in length
 * 88574 88673: gap of 100 bp
 * 88674 93173: contig of 4500 bp in length
 * 93174 93273: gap of 100 bp
 * 93274 101047: contig of 7774 bp in length
 * 101048 101147: gap of 100 bp
 * 101148 111284: contig of 10137 bp in length
 * 111285 11384: gap of 100 bp
 * 11385 130797: contig of 19413 bp in length
 * 130798 130897: gap of 100 bp
 * 130898 152534: contig of 21637 bp in length
 * 152535 152634: gap of 100 bp
 * 152635 182885: contig of 30251 bp in length
 * 182886 182985: gap of 100 bp
 * 182986 184050: contig of 1065 bp in length.

Location/Qualifiers

1..184050
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /chromosome="8"
 /map="8"
 /clone="RP11-507M15"
 /clone_lib="RC11-11 Human Male BAC"
 1..29338
 /note="assembly-fragment"
 clone_end:SP6
 vector_side:left"
 29439..29457
 misc_feature
 /note="assembly-fragment"
 29558..30186
 misc_feature
 /note="assembly-fragment"
 30287..31083
 misc_feature
 /note="assembly-fragment"
 31184..31978
 misc_feature
 /note="assembly-fragment"
 32079..32748
 misc_feature
 /note="assembly-fragment"
 32849..34030
 misc_feature
 /note="assembly-fragment"
 34131..35157
 misc_feature
 /note="assembly-fragment"
 35258..72407
 misc_feature
 /note="assembly-fragment"
 72508..75697
 misc_feature
 /note="assembly-fragment"
 75798..78108
 misc_feature
 /note="assembly-fragment"
 78209..82150
 misc_feature
 /note="assembly-fragment"
 82251..88573
 misc_feature
 /note="assembly-fragment"
 88674..93173
 misc_feature
 /note="assembly-fragment"
 93274..101047
 misc_feature
 /note="assembly-fragment"
 101148..111284
 misc_feature
 /note="assembly-fragment"
 111385..130797
 misc_feature
 /note="assembly-fragment"
 130898..152534
 misc_feature
 /note="assembly-fragment"

Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence.submissions@genome.wi.mit.edu
 Project Information
 Center Project name: L11344
 Center clone name: 507_M_15
 Sequencing vector: Plasmid; n/a; 100% of reads
 Chemistry: Dye-terminator Big Dye; 100% of reads
 Assembly program: Phrap; version 0.960731
 Consensus quality: 175465 bases at least Q40
 Consensus quality: 179683 bases at least Q30
 Consensus quality: 181216 bases at least Q20
 Insert size: 170000; agarose-fp
 Insert size: 182150; sum-of-contigs
 Quality coverage: 7.1 in Q20 bases; agarose-fp
 Quality coverage: 6.7 in Q20 bases; sum-of-contigs
 NOTE: This is a 'working draft' sequence. It currently
 consists of 20 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.
 1 29338: contig of 29338 bp in length
 * 29339 29438: gap of 100 bp
 * 29439 29457: contig of 19 bp in length
 * 29458 29557: gap of 100 bp
 * 29558 30186: contig of 629 bp in length
 * 30187 30286: gap of 100 bp
 * 30287 31083: contig of 797 bp in length
 * 31084 31183: gap of 100 bp
 * 31184 31978: contig of 795 bp in length
 * 31979 32078: gap of 100 bp
 * 32079 32748: contig of 670 bp in length
 * 32749 32848: gap of 100 bp

8132. .0220

	repeat_region	/note="16 copies 3 mer cct 77 conserved"
	repeat_region	16173..16214
	repeat_region	/note="7 copies 6 mer tcctcc 81 conserved"
	repeat_region	17561..17987
	repeat_region	/note="L2 repeat: matches 1231..1706 of consensus"
	repeat_region	18193..18242
	repeat_region	/note="L2 repeat: matches 2702..2750 of consensus"
	repeat_region	18243..18412
	repeat_region	/note="MER5B repeat: matches 1..175 of consensus"
	repeat_region	18413..18473
	repeat_region	/note="L2 repeat: matches 2638..2702 of consensus"
	repeat_region	complement(18442..18745)
	repeat_region	/note="match: STS: Em:HS40462A5; match: STS: Em:Z51609"
	repeat_region	18481..18516
	repeat_region	/note="3 copies 12 mer 94 conserved"
	repeat_region	complement(18698..19137)
	repeat_region	/note="match: GSS: Em:AQ499165"
	repeat_region	19252..19406
	repeat_region	/note="MER9A repeat: matches 11..193 of consensus"
	repeat_region	19642..19730
	repeat_region	/note="MIR repeat: matches 57..145 of consensus"
	repeat_region	20478..20578
	repeat_region	/note="MIR repeat: matches 9..104 of consensus"
Query Match	Best Local Similarity	18.6%; Score 171.6; DB 9; Length 86574;
Matches 299;	Conservative	71.9%; Pred. Mis.9.8e-32;
		Matches 74; Indels 43; Gaps 4
Oy	428 AAGGGGAGGAGGAGAAGGAAGAAAGAGATGATGAGTGCAGATATATCTC	487 /note="16 copies 3 mer cct 77 conserved"
Dd	16224 AGCGAAAAAGGGAGGAGAGGAGAGGAAGATGATGAGATTAAGATTCGCAAAAGGTCTC	16165 /note="7 copies 6 mer tcctcc 81 conserved"
Oy	488 TGGAGAGCGAAAGCCCTGTCTCAACAAAGTATAAAGGTGTGCCCCAGAAAGCAGCCAGCG	547 /note="L2 repeat: matches 1231..1706 of consensus"
Dd	16164 TTGAGAGAGGAGACCCTGT-----CAAGTCAAAAAGCGTTGGCGCCCGCAAGAGCAGCACGCT	16109 /note="L2 repeat: matches 2702..2750 of consensus"
Oy	548 TGGCTTAGAAAAAAAAAGCTGCAAAAAAGAGA-----AGAGCAATATAGAGC	593 /note="MER5B repeat: matches 1..175 of consensus"
Dd	16108 TTGCGAAAAAAGAAAAAGAAAAAGAAAAAAACTGGAAAAAGAGAGAAAGTGAGACC	16049 /note="L2 repeat: matches 2638..2702 of consensus"
Oy	594 CAGCGTTAGACAAGAGAGCCCTGTGAAAAAGGCCAAGCCACAGACCCAGAGCCAGAAAGGCC	653 /note="match: STS: Em:HS40462A5; match: STS: Em:Z51609"
Dd	16048 TAGTGTTAGAGCAAGAGCCCTTGGAAAAATAGGCCAAMCACAATTTCAGGCC-----	15998 /note="3 copies 12 mer 94 conserved"
Oy	654 AGGATTCAAGAANTAGGAGCAGCGCTTGGGGGGCAGCGTGCACAAAGTGGGCTTCCCTG	713 /note="match: GSS: Em:AQ499165"
Dd	15997 -----CAAAGAATATAGAGAGCCATGCTTGGAGGTGATGTCAGAGTGAACCTGGCCCT	15944 /note="MIR repeat: matches 57..145 of consensus"
Oy	714 GGCTGTGCTGCACAGCACAAGGTGCCCTGTGCCAGCCCTCCACCTGTGTCTGAATGACAA	773 /note="MIR repeat: matches 9..104 of consensus"
Dd	15943 GGCTGTGCTGTGGGTGTCAGAGGATGC-----CCCCCTCATGTGTCTGAATATGAC	15894 /note="L2 repeat: matches 1231..1706 of consensus"
Oy	774 AGGGGTGTGGCGGGGCAACATGAGAGCCCTCACCCCACACTGTCCACTTTACAG	829 /note="L2 repeat: matches 2638..2702 of consensus"
Dd	15893 AGGGGTGTGTGGGGGCAACACTAAAGTCTCCACCCCGACTCTTCAGGTGAGG	15838 /note="L2 repeat: matches 1231..1706 of consensus"
RESULT 9	AC108486	135206 bp DNA linear HTG_29-JAN-2002
LOCUS	AC108486	Homo sapiens chromosome 3 clone RP11-114C6, WORKING DRAFT SEQUENCE.
DEFINITION		10 unordered pieces.
ACCESSION	AC108486	
VERSION	AC108486.1	GI:18390285
KEYWORDS	HTG; HTGS_PHAISEI; HTGS_DRAFT; HTGS_FULLTOP; HTGS_ACTIVEFIN.	
SOURCE	Homo sapiens	
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.	
REFERENCE	1 (bases 1 to 135206)	Kaul,R.K., Olson,M.V., Raymond,C. and Haugen,E.D.
TITLE	Direct Submission	
JOURNAL	Unpublished	

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 23, 2003, 14:00:24 ; Search time 310 Seconds

(without alignments)
6712.408 Million cell updates/sec

Title: US-09-844-864-16

Perfect score: 924
Sequence: 1 cagccgcctctctgcgcg.....tttgcgcgcgaagtatg 924

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_101002:*

- 1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
- 2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
- 3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
- 4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
- 5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
- 6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
- 7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
- 8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
- 9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
- 10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
- 11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
- 12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
- 13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
- 14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
- 15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
- 16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
- 17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
- 18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
- 19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
- 20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
- 21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
- 22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
- 23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
- 24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	858	92.9	1048	21	AAAT2432
2	342.8	37.1	1019	21	AAD00296
3	337	36.5	423	23	AAD02440
4	171.6	18.6	86574	24	ABK85560
5	131.6	14.2	1481	21	ABK00297
6	72.4	7.8	416	22	AAL01804
7	70.2	7.6	416	23	ABL97097
8	70.2	7.6	49999	20	AAZ23891
9	70.2	7.6	49999	20	AAZ23896

10	69.2	7.5	33923	22	AAK67071	Human immune/haema
11	68.8	7.4	341	23	AAS90676	DNA encoding novel
12	67.8	7.3	512	24	ABO54697	Human ovarian anti
13	66.4	7.2	291	23	AAS68926	DNA encoding novel
14	65.8	7.1	1680	23	AAS68538	DNA encoding novel
15	65.2	7.1	8165	22	AAK82672	Human immune/haema
16	64	6.9	510	23	AAS69539	DNA encoding novel
17	64	6.9	510	23	AAS71141	DNA encoding novel
18	64	6.9	510	23	AAS90687	DNA encoding novel
19	63.8	6.9	591	23	AAS70521	DNA encoding novel
20	63.8	6.9	591	23	AAS70706	DNA encoding novel
21	63.8	6.9	591	23	AAS90721	DNA encoding novel
22	63.8	6.9	1416	23	AAS67163	DNA encoding novel
23	63.4	6.9	575	22	ABA50472	Human breast cell
24	63.4	6.9	575	22	ABA68422	DNA encoding novel
25	63.4	6.9	575	22	ABA55413	Human foetal liver
26	63.4	6.9	575	22	AAK16793	Human brain expres
27	63.4	6.9	575	22	AAK42567	Human bone marrow
28	63.4	6.9	575	22	AAI23315	Probe #13248 for g
29	63.4	6.9	575	22	AAI18636	Probe #17322 used
30	63.4	6.9	575	22	AAI08956	Probe #8947 used t
31	63.4	6.9	575	22	ABSI6615	Human genome-deriv
32	63.4	6.9	1969	22	ABA45341	Human breast cell
33	63.4	6.9	1969	22	ABA55830	Human foetal liver
34	63.4	6.9	1969	22	ABA25506	Probe #3972 for ge
35	63.4	6.9	1969	22	AAK04048	Human brain expres
36	63.4	6.9	1969	22	AAK29533	Human bone marrow
37	63.4	6.9	1969	22	AAI14105	Probe #4038 for ge
38	63.4	6.9	1969	22	AAI35486	Probe #4172 used t
39	63.4	6.9	1969	22	AAI03958	Probe #3949 used t
40	63.4	6.9	1969	24	ABSO4084	Human genome-deriv
41	63.4	6.9	51259	18	AAK83007	Partial mouse WRN
42	63.2	6.8	14798	24	ABL33033	Human immune syste
43	63.2	6.8	1080	24	AB199537	Mouse ischemic co
44	62.6	6.8	654	23	AAS68925	DNA encoding novel
45	62.4	6.8	575	22	ABA63148	Human foetal liver

ALIGNMENTS

RESULT 1
ID AAAT2432 standard; CDNA; 1048 BP.
XX
AC AAAT2432:
XX
DT 19-DEC-2000 (first entry)
XX
DE Human nucleic acid-binding protein NuBP-51 cDNA.
XX
KW Human nucleic acid-binding protein; NuBP; agonist; antagonist; EST;
KW expressed sequence tag; drug screening; recombinant expression; antibody;
KW reproductive disorder; infertility; immunological disorder;
KW neurological disorder; cell proliferative disorder; cancer; tumour; ss.
XX
OS Homo sapiens
XX
CN MO20004900-A2.
XX
PD 03-AUG-2000.
XX
PF 28-JAN-2000; 2000WO-US02237.
XX
PR 29-JAN-1999; 99US-0117904.
PR 29-JAN-1999; 99US-0117905.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Tang YT, Lal P, Hillman JL, Yue H, Azimzal Y, Lu AMD, Baughn MR;
PI Tran B, Shih LL, Au-Young JL;
XX
DR WPI; 2000-499332/44.

DR P-PSDB; AAB21047.

PT Novel nucleic acid binding proteins, used to identify agonists and
PT antagonists of them, for the treatment of reproductive, immunological,
PT neurological and cell proliferative disorders including cancer -

PS Claim 4; Page 178; 180pp; English.

CC Sequence AAA72382-A72436 represent cDNAs encoding novel human nucleic
CC acid-binding proteins (NuABPs; AAB20997-B21051). These cDNAs were
CC produced by extension from an appropriate EST (expressed sequence
CC tag) using primers designed using the EST. The invention also relates
CC to expression constructs, host cells and transgenic organisms comprising
CC a human NuABP nucleic acid, recombinant production of the human NuABPs,
CC and antibodies against the human NuABPs, and also to methods of
CC screening modulators of human NuABP activity or expression. The human
CC NuABPs, and their agonists and antagonists are used to treat diseases
CC associated with overexpression or underexpression of functional NuABPs.
CC Human NuABP proteins and nucleotides, and NuABP agonists and antagonists
CC can be used to diagnose, treat and prevent reproductive, immunological,
CC neurological and cell proliferative disorders. Reproductive disorders
CC that may be treated using compositions of the invention include
CC infertility, endometriosis, disruptions of the menstrual cycle and
CC disruptions of spermatogenesis. Immunological disorders that may be
CC treated include AIDS, allergies, and autoimmune disorders such as
CC multiple sclerosis, rheumatoid arthritis, diabetes and systemic lupus
CC erythematosus. Neurological disorders that may be treated include
CC epilepsy, neurodegenerative conditions such as Alzheimer's disease and
CC Parkinson's disease, prion diseases such as Creutzfeldt-Jakob disease,
CC and mental disorders such as schizophrenia. Cell proliferative disorders
CC that may be treated include a wide variety of cancers, and also
CC arteriosclerosis, atherosclerosis, cirrhosis and psoriasis.

Sequence 1048 BP; 256 A; 289 C; 338 G; 165 T; 0 other;

Query Match	92.98;	Score 858;	DB 21;	Length 1048;
-------------	--------	------------	--------	--------------

```
Matches 858; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY	1	CAGCCGCTTCTGTCGCCCGGAGCCATGATTCGATGAGCGGCATATACAGAGGAGAAAG	60
Db	178	CAGCCGCTTCTGTCGCCCGGAGCATGATTCGATGAGCGGCATATACAGAGGAGAAAG	237
QY	61	GCAGTACAGACCGTGCCTCGGCGGCGGGAGCTCAGTCAGAGAGAGGCGACTTGGACCTTC	120
Db	238	GCAGTACAGACCGTGCCTCGGCGGCGGGAGCTCAGTCAGTCAGAGAGAGGCGCGACTTGGACCTTC	297
QY	121	AGACCCCGACTGAGAGGGGAACAGAGCTGCAGGCTGTGTCTTCATACGATTTGCTTGGGG	180
Db	298	AGACCCCGACTGAGAGGGGAACAGAGCTGCAGGCTGTGTCTTCATACGATTTGCTTGGGG	357
QY	181	GGAAGAGCCAAAGAGAGATGCATCGGTGAGATCCGCCCCAGCAAAACAGAGAGAC	240
Db	358	GGAAGAGCCAAAGAGAGATGCATCGGTGAGATCCGCCCCAGCAAAACAGAGAGAC	417
QY	241	AGAAGATGCAGCCGGTCACCATTTGCTCACTCCAGGCTCAGTCGTCGCCCATGTCYC	300
Db	418	AGAAGATGCAGCCGGTCACCATTTGCTCACTCCAGGCTCAGTCGTCGCCCATGTCYC	477
QY	301	ATGGTAGAGTGCACACTTTCCTCCCGAGTTTCCTTCAGAGTCGCGGGCGGGGTAGAGACC	366
Db	478	ATGGTAGAGTGCACACTTTCCTCCCGAGTTTCCTTCAGAGTCGCGGGCGGGGTAGAGACC	537
QY	361	GTTGTCCTCAGTGGCCAGGAAGCTTATGAAACATCAGACCTAACCTGGAGAGAGAGAG	420
Db	538	GTTGTCCTCAGTGGCCAGGAAGCTTATGAAACATCAGACCTAACCTGGAGAGAGAGAG	597
QY	421	GAAAGAAAGAGGGAGAGAGAGGAAGAAGGAAGAAGTATGATGAGATGAGAGATGCAGAT	480
Db	598	GAAAGAAAGAGGGAGAGAGAGGAAGAAGGAAGAAGTATGATGAGATGAGAGATGCAGAT	655
QY	481	ATATCTCTGAGAGACCAAGCCCTGTCCAACAGTCAAAAGCTGTGTCGCCCAGAAACAG	540

Db	658	ATATCTCTGGAGGAGCAAAAGCCCTGTCAACCAAGTCAAAAAGGCTGGTGCCTCCAGAGACAG	712
Qy	541	CGGAGCGTGCTAAGAAAAAAAAGCTTGGAAAAAGAGAGAGGAATTAAGAGCCACAGTT	600
Db	718	GCGAGCGTGCTAAGAAAAAAAAGCTTGGAAAAAGAAAGAAATTAAGAGCCACAGTT	777
Qy	601	AGACACAAGAGCCCTGTGAAAAAAGGCCAAGGCCACAGCCAGAGCCCAAGAGCAGGATTG	660
Db	778	AGACACAAGAGCCCTGTGAAAAAAGGCCAAGGCCACAGCCAGAGCCCAAGAGCAGGATTG	837
Qy	661	AAGAAATGAGAGGACGACAGCCCTTGGGGGGGACAGGTGCAAAAGTGGGCGCTTCCCTGGGCTGTG	720
Db	838	AAGAAATGAGAGGACGACAGCCCTTGGGGGGGACAGGTGCAAAAGTGGGCGCTTCCCTGGGCTGTG	897
Qy	721	CTGACGACACAGGCTGCCCTGTCTCAGCCCTCCACTGTGTCTGAATGCAACAGGGGTG	780
Db	898	CTGACGACACAGGCTGCCCTGTCTCAGCCCTCCACTGTGTCTGAATGCAACAGGGGTG	957
Qy	781	TTGGGGGGGGCAACATGAGAGCCCTCAACCCCAACCTCCACTTTCAGAGGGGCCCCAGT	840
Db	958	TTGGGGGGGGCAACATGAGAGCCCTCAACCCCAACCTCCACTTTCAGAGGGGCCCCAGT	101
Qy	841	GAAAGAGCCCACTCTGGG	858
Db	1018	GAAAGAGCCCACTCTGGG	1035

RESULT 2

AC AAD00296;

DT 05-SEP-2000 (first entry)

DE Mouse oocyte-specific O1-236 cDNA clone

KM Oocyte-specific; ovary; O1-236; mouse; Npm2; gynaecological; treatment;
 KM nucleoplasmic; cell proliferative disorder; cell degenerative disorder;
 KM contraceptive; ovulation; signalling pathway; human infertility; cancer;
 KM screen; modulator; ss.

Key	Location/Qualifiers
PH	156..779
PT	/*tag= a
PT	/product= "Mouse oocyte-specific protein, O1-236"
PT	/note= "Identical to mouse Npm2 protein with the exception of one residue"
FT	

PN W0200024755-A1

PD 04-MAY-2000.

PF 28-OCT-1999; 99WO-US25209.

PR 28-OCT-1998; 98US-0106020.

PA (BAYU) BAYLOR COLLEGE MEDICINE.

PI Matzuk MM, Wang P;

DR WP1; 2000-350684/30.

XX
XX

PT useful for evaluating potential contraceptives to block ovulation in a

PS Claim 22; Fig 5; 54pp; English.

CC The present sequence is the mouse oocyte-specific gene Npm2, complete

CC cDNA clone O1-236. It is derived from mouse 2-cell embryo cDNA library
 CC and expressed in the oocytes of intermediate size type 3a follicles and
 CC all type 3b follicles. This clone is used to screen and identify the
 CC mouse Nrm2 gene. It is homologous to Xenopus laevis nucleoplasm (Xnpn2)
 CC expressed exclusively in eggs. It provides in vitro and in vivo reagents
 CC for studying ovarian development and function. This sequence has
 CC gynecological and contraceptive activity. Agents which modulate O1-180,
 CC O1-184 and O1-236 may be used to treat cell proliferative or degenerative
 CC disorders, associated with abnormal expression of these ovary specific
 CC genes. This ovary-specific sequence can be used as reagents to evaluate
 CC potential contraceptives, to block ovulation in a reversible manner.
 CC It is also used to screen for genetic mutations in signalling pathways,
 CC that are associated with some forms of human infertility or
 CC gynecological cancers.
 CC
 XX
 XX

SO Sequence 1019 BP; 334 A; 243 C; 265 G; 177 T; 0 other;

Query Match 37.1%; Score 342.8; DB 21; Length 1019;

Best Local Similarity 66.4%; Pred. No. 1.5e-72;

Matches 603; Conservative 0; Mismatches 262; Indels 43; Gaps 6;

OY 1 CACCCCGCTTCTCTGCGGAGCATGATCTCAGTAGCGCCAGTAGACGAGAGAAAG 60
 DB 132 CACCAAGCGCGCTTAAATCGACATGAGTCGCGACAGCAGCAGCGCTGACCGAAACC 191
 OY 61 GCAGTAGCAGACCGCTGCTGGGGCTGCGAGTCAGTAGAGAGAGCGGAGCTTGACCTTC 120
 DB 192 ACAGCAAAAACATGCTGGGGTAGTGAATCAATCAGAAACAGAGACTTGACCTTT 251
 OY 121 AGACCCGACTGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
 DB 252 AGAGGCGCAAGGCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 311
 OY 181 GAGAAAGCCAAAG 240
 DB 312 GAGAAAGCCAAAG 359
 OY 241 AAGAGATCAGAGCGGCTCAGATTCGCTCAGTCCAGGCTCAGTCCCTCCAGTCTCC 300
 DB 360 GCGAGAGAAACCAATCATCTATCTGAGTGAAGGAGTCACTGCTGCTGCTGCTCACT 419
 OY 301 ATGCTGAG 360
 DB 420 GTGTCAAGGTATAGAGTTTCTCTCCAGTAACTTTTGGCTGAGAGAGAGAGAGAG 479
 OY 361 GTGTCTCTCAGTGGCAG 420
 DB 480 GTGTCTCTCAGTGGCAG 539
 OY 421 GAAGAAAG 480
 DB 540 GAA-----GAT 587
 OY 481 ATATCTCTGAG 540
 DB 588 ATATCTCTGAG 644
 OY 541 GCGAGCGCTGCTTAAAGAAAAAGCTGGAAGAAAGAGAGAGAGAGAGAGAGAGAGAG 594
 DB 645 ATGACCATAGCAAGAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 704
 OY 595 AGCGTTAG 654
 DB 705 AGCCCTCAGGAG 764
 OY 655 GATTTCAG 714
 DB 765 GTGACCAAGAAATGA--CTCATCTTACATCTTCTGCTCAGAGAGAGAGAGAGAGAG 822
 OY 715 GCTGTGCTGACAGGAG 774
 DB 823 GCTGTGTT-----TTGTTGTCAGAGTGTCCAGGCGCCACACCCAGTCTGTAATGA 874

OY 775 GGGGTGTGGGGGAGCAATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 834
 DB 875 AGGTGTGTGGGTGTAACTGTAACCTGTAAACCCAGAGAGAGAGAGAGAGAGAGAG 934
 OY 835 CCCACTGAAG 894
 DB 935 AGCCCGCAGCAAGTTCGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 994
 OY 895 AAAAAAAA 902
 DB 995 AAAAAAAA 1002

RESULT 3

AD23440
 ID AAD23440 standard; cDNA: 423 BP.

AC AAD23440;

DT 26-FEB-2002 (first entry)

DE Human lung tumour-specific 54589.1 cDNA.

XX Human lung tumour protein; immunostimulant; cytostatic; gene therapy;
 KW antisense-therapy; vaccine; immune response; lung cancer; 54589.1; 55.

XX Homo sapiens.

PN WO200172295-A2.

PD 04-OCT-2001.

PE 28-MAR-2001; 2001MO-US09991.

PR 29-MAR-2000; 2000US-0538037.

PR 05-JUN-2000; 2000US-0588937.

PR 18-AUG-2000; 2000US-0640878.

PR 22-SEP-2000; 2000US-234517P.

PR 01-NOV-2000; 2000US-0704512.

PR 14-DEC-2000; 2000US-0738973.

XX (CORI-) CORIXA CORP.

PA Reed SG, Lodes MJ, Mohamath R, Secrist H, Benson DR, Indrias CY;
 PI Henderson RA, Fling SP, Algate PA, Elliot M, Kannon J, Kalos MD;
 PT WPI: 2001-639201/73.

PT New human lung-specific polynucleotides and polypeptides for the
 PT diagnosis and treatment of disease e.g. lung cancer -
 XX Claim 1; Page 302; 378bp; English.

XX The invention relates to isolated lung tumour-specific proteins and
 CC their corresponding cDNA molecules. Lung tumour-specific proteins and
 CC their antigen-presenting cells are useful for stimulating and/or
 CC expanding T cells specific for a tumour protein, and for inhibiting
 CC the development of cancer. The invention also relates to a composition
 CC useful for stimulating an immune response, and for treating cancer. The
 CC lung tumour specific oligonucleotide is useful in gene therapy and for
 CC diagnosis, detection and treatment of lung cancer. The present sequence
 CC is a cDNA encoding human lung tumour-specific protein.
 CC
 XX

SO Sequence 423 BP; 84 A; 138 C; 125 G; 76 T; 0 other;

Query Match 36.5%; Score 337; DB 23; Length 423;

Best Local Similarity 100.0%; Pred. No. 2.7e-71;

Matches 337; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 CAGCCCGCTTCTCTGCGGAGCATGATCTCAGTAGCGCCAGTAGAGAGAGAGAGAG 60
 DB 87 CAGCCCGCTTCTCTGCGGAGCATGATCTCAGTAGCGCCAGTAGAGAGAGAGAGAG 146

Chromosome	Gene	Transcript	Feature	Location/Qualifiers
KW	chromosome 14	human chromosome 8p21: centriceptive; gynaecology; cancer		
KW	cell proliferative disorder; cell degenerative disorder; ovulation;			
KW	modulator; human infertility; signalling pathway; screen; treatment; ds.			
OS	Mus sp.			
XX	Key			
XX	5'UTR			
FT				Location/Qualifiers
FT				1..150
FT				/*tag= a
FT	exon			151..199
FT				/*tag= b
FT				/number= 1
FT	Intron			200..230
FT				/*tag= c
FT				/number= 1
FT				215
FT	misc_feature			/*tag= d
FT				/note= "Corresponds to 314 missing nucleotides of
FT				intron 1 between bases 214 and 216"
FT	exon			231..394
FT				/*tag= e
FT				/number= 2
FT	CDS			337..1177
FT				/*tag= f
FT				/product= "Mouse ovary-specific Npm2 protein"
FT				/note= "Coding region is interrupted by 8 introns"
FT	Intron			395..425
FT				/*tag= g
FT				/number= 2
FT	misc_feature			410
FT				/*tag= h
FT				/note= "Corresponds to 105 missing nucleotides of
FT				intron 2 between bases 409 and 411"
FT	exon			426..511
FT				/*tag= i
FT				/number= 3
FT	misc_feature			473
FT				/*tag= j
FT				/note= "The 'T' is replaced with 'G' in the cDNA"
FT	Intron			512..542
FT				/*tag= k
FT				/number= 3
FT	misc_feature			527
FT				/*tag= l
FT				/note= "Corresponds to 63 missing nucleotides of
FT				intron 3 between bases 526 and 528"
FT	exon			543..656
FT				/*tag= m
FT				/number= 4
FT	Intron			657..687
FT				/*tag= n
FT				/number= 4
FT	misc_feature			672
FT				/*tag= o
FT				/note= "Corresponds to 2771 (2.77kb) missing nucleotides
FT				of intron 4 between bases 671 and 673"
FT	exon			688..781
FT				/*tag= p
FT				/number= 5
FT	Intron			782..812
FT				/*tag= q
FT				/number= 5
FT	misc_feature			797
FT				/*tag= r
FT				/note= "Corresponds to 1321 (1.32kb) missing nucleotides
FT				of intron 5 between bases 796 and 798"
FT	exon			813..964
FT				/*tag= s
FT				/number= 6
FT	Intron			965..995
FT				/*tag= t
FT				/number= 6
FT	misc_feature			980

FT	/tag= u
ET	/note= "Corresponds to 157 missing nucleotides of
ET	intron 6 between bases 979 and 981"
ET	996..1036
ET	/tag= v
ET	/number= 7
ET	1037..1067
ET	/tag= w
ET	/number= 7
ET	1052
ET	/tag= x
ET	/note= "Corresponds to 471 missing nucleotides of
ET	intron 7 between bases 1051 and 1053"
ET	1068..1101
ET	/tag= y
ET	/number= 8
ET	1102..1132
ET	/tag= z
ET	/number= 8
ET	1117
ET	/tag= aa
ET	/note= "Corresponds to 63 missing nucleotides of
ET	intron 8 between bases 1116 and 1118"
ET	1133..1381
ET	/tag= ab
ET	/number= 9
ET	1229
ET	/tag= ac
ET	/note= "The 'C' is replaced with 'T' in the cDNA"
ET	1361..1366
ET	/tag= ad
ET	1382..1481
ET	/tag= ae
ET	
XX	WO200024755-A1.
PD	04-MAY-2000.
PE	28-OCT-1999; 99WO-US25209.
PR	28-OCT-1998; 98US-0106020.
XX	(BAYU) BAYLOR COLLEGE MEDICINE.
PA	
PI	Matzuk MM, Wang P;
PI	
DR	WPI: 2000-350684/30.
DR	P-PsDB: AAY70951.
PT	O1-180, O1-184 and O1-236 polypeptides and nucleic acids encoding them,
PT	useful for evaluating potential contraceptives to block ovulation in a
PT	reversible manner -
PS	Example 5; Fig 13; 54pp; English.
XX	
CC	The present sequence is the mouse ovary-specific Npm2 gene, mapped to
CC	the middle of mouse chromosome 14. It shows linkage to D14M132, this
CC	region being syntenic to human chromosome 8p21. The clone O1-236 cDNA is
CC	used to screen and identify Npm2 gene. It is the mammalian ortholog of
CC	Xenopus laevis nucleoplasmal (Xnpm2) expressed exclusively in the eggs.
CC	It provides in vitro and in vivo reagents for studying ovarian
CC	development and function. This sequence has gynaecological and
CC	contraceptive activity. Agents which modulate O1-180, O1-184 and O1-236
CC	may be used to treat cell proliferative or degenerative disorders.
CC	associated with abnormal expression of these ovary specific genes. This
CC	ovary-specific sequence can be used as reagents to evaluate potential
CC	contraceptives, to block ovulation in a reversible manner. It is also
CC	used to screen for genetic mutations in signalling pathways, that is
CC	associated with some form of human infertility or gynaecological cancer
XX	
SQ	Sequence 1481 BP; 407 A; 379 C; 391 G; 296 T; 8 other;

Query Match 14.2%; Score 131.6; DB 21; Length 1481;

PR	26-JUL-2000	2000US-0220963
PR	26-JUL-2000	2000US-0220964
PR	14-AUG-2000	2000US-0224518
PR	14-AUG-2000	2000US-0224519
PR	14-AUG-2000	2000US-0225213
PR	14-AUG-2000	2000US-0225214
PR	14-AUG-2000	2000US-0225266
PR	14-AUG-2000	2000US-0225267
PR	14-AUG-2000	2000US-0225268
PR	14-AUG-2000	2000US-0225270
PR	14-AUG-2000	2000US-0225447
PR	14-AUG-2000	2000US-0225757
PR	14-AUG-2000	2000US-0225758
PR	14-AUG-2000	2000US-0225759
PR	18-AUG-2000	2000US-0226279
PR	22-AUG-2000	2000US-0226681
PR	22-AUG-2000	2000US-0226686
PR	22-AUG-2000	2000US-0227182
PR	23-AUG-2000	2000US-0227709
PR	30-AUG-2000	2000US-0228924
PR	01-SEP-2000	2000US-0229287
PR	01-SEP-2000	2000US-0229344
PR	01-SEP-2000	2000US-0229345
PR	05-SEP-2000	2000US-0229349
PR	05-SEP-2000	2000US-0229513
PR	06-SEP-2000	2000US-0230437
PR	06-SEP-2000	2000US-0230438
PR	08-SEP-2000	2000US-0231243
PR	08-SEP-2000	2000US-0231244
PR	08-SEP-2000	2000US-0231413
PR	08-SEP-2000	2000US-0231414
PR	08-SEP-2000	2000US-0232080
PR	08-SEP-2000	2000US-0232081
PR	12-SEP-2000	2000US-0231968
PR	14-SEP-2000	2000US-0233397
PR	14-SEP-2000	2000US-0233398
PR	14-SEP-2000	2000US-0233399
PR	14-SEP-2000	2000US-0234000
PR	14-SEP-2000	2000US-0234401
PR	14-SEP-2000	2000US-0234063
PR	14-SEP-2000	2000US-0234064
PR	14-SEP-2000	2000US-0234065
PR	27-SEP-2000	2000US-0235836
PR	29-SEP-2000	2000US-0236327
PR	29-SEP-2000	2000US-0236367
PR	29-SEP-2000	2000US-0236368
PR	29-SEP-2000	2000US-0236369
PR	29-SEP-2000	2000US-0236370
PR	02-OCT-2000	2000US-0236802
PR	02-OCT-2000	2000US-0237032
PR	02-OCT-2000	2000US-0237038
PR	02-OCT-2000	2000US-0237039
PR	02-OCT-2000	2000US-0237040
PR	13-OCT-2000	2000US-0239353
PR	13-OCT-2000	2000US-0239357
PR	20-OCT-2000	2000US-0241920
PR	20-OCT-2000	2000US-0241921
PR	20-OCT-2000	2000US-0241785
PR	20-OCT-2000	2000US-0241786
PR	20-OCT-2000	2000US-0241787
PR	20-OCT-2000	2000US-0241808
PR	20-OCT-2000	2000US-0241809
PR	20-OCT-2000	2000US-0244617
PR	01-NOV-2000	2000US-0244618
PR	08-NOV-2000	2000US-0246474

PR 08-NOV-2000; 2000US-0246475.
 PR 08-NOV-2000; 2000US-0246476.
 PR 08-NOV-2000; 2000US-0246477.
 PR 08-NOV-2000; 2000US-0246478.
 PR 08-NOV-2000; 2000US-0246479.
 PR 08-NOV-2000; 2000US-0246523.
 PR 08-NOV-2000; 2000US-0246524.
 PR 08-NOV-2000; 2000US-0246525.
 PR 08-NOV-2000; 2000US-0246526.
 PR 08-NOV-2000; 2000US-0246527.
 PR 08-NOV-2000; 2000US-0246528.
 PR 08-NOV-2000; 2000US-0246532.
 PR 08-NOV-2000; 2000US-0246609.
 PR 08-NOV-2000; 2000US-0246610.
 PR 08-NOV-2000; 2000US-0246611.
 PR 08-NOV-2000; 2000US-0246613.
 PR 17-NOV-2000; 2000US-0249207.
 PR 17-NOV-2000; 2000US-0249208.
 PR 17-NOV-2000; 2000US-0249209.
 PR 17-NOV-2000; 2000US-0249210.
 PR 17-NOV-2000; 2000US-0249211.
 PR 17-NOV-2000; 2000US-0249212.
 PR 17-NOV-2000; 2000US-0249213.
 PR 17-NOV-2000; 2000US-0249214.
 PR 17-NOV-2000; 2000US-0249215.
 PR 17-NOV-2000; 2000US-0249216.
 PR 17-NOV-2000; 2000US-0249217.
 PR 17-NOV-2000; 2000US-0249218.
 PR 17-NOV-2000; 2000US-0249244.
 PR 17-NOV-2000; 2000US-0249245.
 PR 17-NOV-2000; 2000US-0249264.
 PR 17-NOV-2000; 2000US-0249265.
 PR 17-NOV-2000; 2000US-0249267.
 PR 17-NOV-2000; 2000US-0249299.
 PR 17-NOV-2000; 2000US-0249300.
 PR 01-DEC-2000; 2000US-0250160.
 PR 01-DEC-2000; 2000US-0250391.
 PR 05-DEC-2000; 2000US-0251030.
 PR 05-DEC-2000; 2000US-0251988.
 PR 05-DEC-2000; 2000US-0256719.
 PR 06-DEC-2000; 2000US-0251479.
 PR 08-DEC-2000; 2000US-0251856.
 PR 08-DEC-2000; 2000US-0251868.
 PR 08-DEC-2000; 2000US-0251869.
 PR 08-DEC-2000; 2000US-0251989.
 PR 08-DEC-2000; 2000US-0251990.
 PR 11-DEC-2000; 2000US-0254097.
 PR 05-JAN-2001; 2001US-0259678.
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Rosen CA, Barash SC, Ruben SM;
 XX WPI: 2001-465570/50.
 DR P-PSDB: AAM95834.
 XX
 PT Isolated nucleic acid molecule encoding a reproductive system antigen
 PT is used in preventing, treating or ameliorating a medical condition -
 XX
 PS Claim 1: SEQ ID NO 1805; 1297bp + Sequence Listing: English.
 XX
 CC The present invention provides the protein and coding sequences of a
 CC number of human reproductive system related antigens. These can be used
 CC in the prevention and treatment of reproductive system disorders,
 CC including cancer. The present sequence is a coding sequence of the
 CC invention.
 XX
 SQ Sequence 416 BP; 96 A; 105 C; 130 G; 80 T; 5 other;

Query Match 7.8%; Score 72.4; DB 22; Length 416;
 Best Local Similarity 50.0%; Pred.No. 8.2e-08;
 Matches 203; Conservative 1; Mismatches 196; Indels 6; Gaps 1;

QY 51 GGAGGAAAGGACGACGACCGCTCTGCGGGCTCGAGCTCAGTCAGAGAGCGGAC 110

Db 2 GCACGAGAGAGGTGTACTGGGCGGTAAGAGGCTGTGAGCTCTCCGCCACACCGCTC 61
 QY 111 TTGGACCTTCAGACCCAGCTGGAGGGAGAGAGAGCTGACGCTGTGCTTATACGAT 170
 Db 62 CTTACCTTTTAAAGTTAGAGAGAGAGATGATGCGGASACGTGTGCACTAACATGCT 121
 QY 171 TTGCTTGGGGAGAAAGCCAAAGAGAGATGATTCGCTGTGAGATCCTGCCACCAA 230
 Db 122 CTGCTCAGCGAGAGAGACCCAAAGACGATGATGTGTGAAGTTGTGCCCGGA---- 177
 QY 231 CCAGGAGACAAAGATGACGACCGGTACCATTTGCTCCTCAGCCCTCACTCTCC 290
 Db 178 -ACCATGACCATCAGAGATGATTCGACCTGTGGCCAACTCAACTGCTGCCAAC 235
 QY 291 CATGCTTCATGATGAGATGATGACGCTTTTCCCCAGTTACTTTTCACTGCGGGCTG 350
 Db 236 CATGCTCAGTCTGTGATGATCTTCAGCTCCAAACCACTTCAACTTCCGCTGAAGTCGG 295
 QY 351 CTCAGACCCGCTGTTCTCAGTGGCCAGAGCGTTATGAAGCATTCAGACCTAACCTG 410
 Db 296 TTCTGGCCCTGTGCGGATCANTGGCGGCACACGATTTGTACATGACCAATATGTTTC 355
 QY 411 GGAGGAGAGAGAAAGAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 456
 Db 356 TGAG 401

RESULT 7
 ABL97097
 ID ABL97097 standard; cDNA; 416 BP.
 XX
 AC ABL97097;
 XX
 DT 21-JUN-2002 (first entry)
 XX
 DE Human testicular antigen encoding cDNA SEQ ID NO: 765.
 XX
 KW Human: testicular antigen; testes; cancer; metastasis; immune disorder;
 KW reproductive system disorder; urinary system disorder; gene therapy;
 KW gastrointestinal disease; infectious; cytostatic; gene; ss.
 OS Homo sapiens.
 XX
 PN WO200155317-A2.
 XX
 PD 02-AUG-2001.
 XX
 PF 17-JAN-2001; 2001WO-US01329.
 XX
 PR 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.

PR	08-NOV-2000	2000US-0246609.
PR	08-NOV-2000	2000US-0246610.
PR	08-NOV-2000	2000US-0246611.
PR	08-NOV-2000	2000US-0246613.
PR	17-NOV-2000	2000US-0249207.
PR	17-NOV-2000	2000US-0249208.
PR	17-NOV-2000	2000US-0249209.
PR	17-NOV-2000	2000US-0249210.
PR	17-NOV-2000	2000US-0249211.
PR	17-NOV-2000	2000US-0249212.
PR	17-NOV-2000	2000US-0249213.
PR	17-NOV-2000	2000US-0249214.
PR	17-NOV-2000	2000US-0249215.
PR	17-NOV-2000	2000US-0249216.
PR	17-NOV-2000	2000US-0249217.
PR	17-NOV-2000	2000US-0249218.
PR	17-NOV-2000	2000US-0249219.
PR	17-NOV-2000	2000US-0249245.
PR	17-NOV-2000	2000US-0249264.
PR	17-NOV-2000	2000US-0249265.
PR	17-NOV-2000	2000US-0249287.
PR	17-NOV-2000	2000US-0249299.
PR	17-NOV-2000	2000US-0249299.
PR	01-DEC-2000	2000US-0250391.
PR	05-DEC-2000	2000US-0251030.
PR	05-DEC-2000	2000US-0251988.
PR	05-DEC-2000	2000US-0251988.
PR	06-DEC-2000	2000US-0256719.
PR	08-DEC-2000	2000US-0251479.
PR	08-DEC-2000	2000US-0251856.
PR	08-DEC-2000	2000US-0251868.
PR	08-DEC-2000	2000US-0251869.
PR	08-DEC-2000	2000US-0251989.
PR	11-DEC-2000	2000US-0251990.
PR	05-JAN-2001	2000US-0254097.
PR	05-JAN-2001	2001US-0259678.

PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX
PI Rosen CA, Barash SC, Ruben SM;
XX
XX
DR WPI; 2001-483426/52.
XX

XX Disclosure: SEQ ID NO 21863: 3071pp + Sequence Listing: English.

XX DISCLOSURE; SEQ ID NO 21883; 3071pp + Sequence Listing; English.

SO	Sequence	33923	BP:	9925	A:	8140	C:	7946	G:	7912	T:	0	other:
	Query Match	7.58;	Score	69.2;	DB	22;	Length	33923;					
	Best Local Similarity	58.18;	Pred. No.	2e-06;									
	Matches	122;	Conservative	0;	Mismatches	80;	Indels	0;	Gaps	0;			

Oy	377	AGGAACGTTATTAAGAAGCATCAGACTCAAACTGGGAGAGAGAGAGAGAGAACAGACAGCGAGC	436
Dd	22820	AAGAAGAGAGAAAGGAGGAGGACGACGAGAGAGAGACACAAGAGAGAGAGACACAAGAGAGAGAGC	22879
Oy	437	AGGAGAGAAAGAGAGAGAGAGAGATGATGAGGATTGAGCATTCAGATATTATCTCTGGAGAGGC	496
Dd	22880	AGGAGAGAGAGACGACGAGGAGGACGAGAGGACGAGAGAGAGAGAGAGAGAGAGAGAGAGG	22933
Oy	497	AAAGCCCTGTCAAACCAAGTCAAAAAGGCTGTGCCCCAGAAAGCAGGCGAGCGCTGGCTAGCA	556
Dd	22940	AGGAAGGAGGAGAAAGCAGGAGAAAGAGAAAGAGAAAGAGAGAGAGAGAGAGAGAGAGAAAGC	22999
Oy	557	AAAAAAGCTGAAAAAAGAAAGAGAGAGAA	586
Dd	23000	AGGAGAAAGAGAGAAAGAGAGAGAGAGAA	23079

RESULT 11
AAS90676
ID AAS90676 standard; CDNA; 341 BP
vv

AC AAS90676;

DT 13-FEB-2002 (first entry)

DNA encoding novel human diagnostic protein #26480.

Human; chromosome mapping; gene therapy; forensic;
 food supplement; medical imaging; diagnostic; genetic disorder; ss
 xy

OS Homo sapiens

PN WO2001.75067-A2

PD 11-OCT-2001

PF 30-MAR-2001; 2001WO-US08631

PR 31-MAR-2000; 2000US-0540217

XX
XXXX
DT

D
E
C
E
M
B
E
R

1
9
7
0

U
S
A

XX
XX
LETT.
C
C
C
C
C
C
G
C

DR P-PSDB; ABG26489.

PT New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity -

PS Claim 1; SEQ ID NO 26480; 103pp; English

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridization probes, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (I) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. AAS4157-AAS4564 represent novel human diagnostic coding sequences of the invention.

AC AAS68926;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE DNA encoding novel human diagnostic protein #4730.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
 OS Homo sapiens.
 XX
 PN W0200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001MO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 XX 23-AUG-2000; 2000US-0649167.
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI: 2001-639362/73.
 DR P-PSDB: ABG04739.
 XX
 PS New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 XX Claim 1: SEQ ID NO 4730; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 XX
 SO Sequence 291 BP; 129 A; 22 C; 112 G; 28 T; 0 other;
 Query Match 7.2%; Score 66.4; DB 23; Length 291;
 Best Local Similarity 54.0%; Pred. No. 2e-06;
 Matches 136; Conservative 0; Mismatches 116; Indels 0; Gaps 0;
 Oy 363 TGTTCCTCAGTCCGAGAACGTTATGACCATTAACCTTACCTGGAGAGAGAGG 421
 Db 8 TTTGCCATATTTGGCCAGGCTAGTCTCGAATCTTGACCTAGGAGGAGAGAGAGG 67
 Oy 422 AAGAAGAGAGGAG 481
 Db 68 AGGAG 127
 Oy 482 TATCTCTGAGAGCAAAAGCCCTCTCAACAGTCAAAAGCGTGTGCCCCCAGAGCAGG 541
 Db 128 AAG 187

Oy 542 CGAGCTGGCTAAGAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 601
 Db 188 AAGAAG 247
 Oy 602 GAGACAGAGGCC 613
 Db 248 GAGATCAAAATCC 259
 RESULT 14
 AAS68538
 ID AAS68538 standard; cDNA; 1680 BP.
 XX
 AC AAS68538;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE DNA encoding novel human diagnostic protein #4342.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
 OS Homo sapiens.
 XX
 PN W0200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001MO-US08631.
 XX
 PR 31-MAR-2000; 2000US-0540217.
 XX 23-AUG-2000; 2000US-0649167.
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;
 XX
 DR WPI: 2001-639362/73.
 DR P-PSDB: ABG04351.
 XX
 PS New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity -
 XX
 XX Claim 1: SEQ ID NO 4342; 103pp; English.
 XX
 CC The invention relates to isolated polynucleotide (I) and
 CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
 CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
 CC and gene mapping, and in recombinant production of (II). The
 CC polynucleotides are also used in diagnostics as expressed sequence tags
 CC for identifying expressed genes. (I) is useful in gene therapy techniques
 CC to restore normal activity of (II) or to treat disease states involving
 CC (II). (II) is useful for generating antibodies against it, detecting or
 CC quantitating a polypeptide in tissue, as molecular weight markers and as
 CC a food supplement. (II) and its binding partners are useful in medical
 CC imaging of sites expressing (II). (I) and (II) are useful for treating
 CC disorders involving aberrant protein expression or biological activity.
 CC The polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human
 CC diagnostic coding sequences of the invention.
 CC Note: The sequence data for this patent did not appear in the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 XX
 SO Sequence 1680 BP; 574 A; 286 C; 485 G; 335 T; 0 other;
 Query Match 7.1%; Score 65.8; DB 23; Length 1680;

Contact: Robert Strausberg, Ph.D.
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: Ash Alizadeh, John Byrd, M.D., Mike Grever,
M.D., Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution Information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/bdnp/image/image.html
Seq primer: -400p from Gibco

FEATURES High quality sequence stop: 457.
Location/Qualifiers
1. 517
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="NCI-CGAP-CL11"
/issue_type="B-cell, chronic lymphocytic leukemia"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5',
TGTTCACATCTGAAGTGGAGGAGGCGCCATTCCTTTTCTTTT
TGTTCACATCTGAAGTGGAGGAGGCGCCATTCCTTTTCTTTT
T 3']; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT7T3 vector.
Library is normalized, and was constructed by Bento
Soares and M. Fatima Bonaldo."

BASE COUNT 70 a 168 c 121 g 158 t
ORIGIN

Query Match 56.0%; Score 517; DB 9; Length 517;
Best Local Similarity 100.0%; Pred. No. 6.1e-77;
Matches 517; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 367 CTCAGTGGCCAGACCTTATGAAGCATCAGACCTTAACCTGGAGAGAGAGAGAGAA 426
DB 517 CTCAGTGGCCAGACCTTATGAAGCATCAGACCTTAACCTGGAGAGAGAGAGAA 458
QY 427 GAAGGGGAGAGAGAGAGAGAGAGAGATGATGAGATGAGATGATGATGATGATGAT 486
DB 457 GAAGGGGAGAGAGAGAGAGAGAGAGATGATGAGATGAGATGATGATGATGATGAT 398
QY 487 CTGAGAGCAAAAGCCTGTCAAAAGTCAAAAGCTGTGCTCCCAAGAGAGAGAGAG 546
DB 397 CTGAGAGCAAAAGCCTGTCAAAAGTCAAAAGCTGTGCTCCCAAGAGAGAGAGAG 338
QY 547 GTGGCTAG 606
DB 337 GTGGCTAG 278
QY 607 AAGAGCCCTGTGAAAAAGGCAAGCCAGAGCCAGAGCCAGAGATTCAGAGAA 666
DB 277 AAGAGCCCTGTGAAAAAGGCAAGCCAGAGCCAGAGCCAGAGATTCAGAGAA 218
QY 667 TGAG 726
DB 217 TGAG 158
QY 727 GCACAG 786
DB 157 GCACAG 98
QY 787 GGGCAACATGAG 846
DB 97 GGGCAACATGAG 38
QY 847 CCCACCTCGGGGTCAACATTAAGTGTGCTGTCAGG 883
DB 37 CCCACCTCGGGGTCAACATTAAGTGTGCTGTCAGG 1

RESULT 2
A1016313/c 455 bp mRNA linear EST 27-AUG-1998
LOCUS A1016313
DEFINITION ot/8a10.s1 Soares_total_fetus_Nb2Hr8_9w Homo sapiens cDNA clone
IMAGE:162874 3' similar to contains element MER22 repetitive
element 3' mRNA sequence.
ACCESSION A1016313
VERSION A1016313.1 GI:3230649
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 455)
AUTHORS NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
Insert Length: 1368 Std Error: 0.00
Seq primer: -40m13 fwd. 5' from Amersham
High quality sequence stop: 447.
Location/Qualifiers

FEATURES
source 1. 455
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Soares_total_fetus_Nb2Hr8_9w"
/dev_stage="8-9 weeks"
/lab_host="DH10B"
/note="Vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker. Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from pooled 8-9 week
(total) fetus material with a Not I - oligo(dT) primer [5',
TGTTCACATCTGAAGTGGAGGAGGCGCCATTCCTTTTCTTTT
TGTTCACATCTGAAGTGGAGGAGGCGCCATTCCTTTTCTTTT
T 3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 62 a 143 c 110 g 140 t
ORIGIN

Query Match 49.2%; Score 455; DB 9; Length 455;
Best Local Similarity 100.0%; Pred. No. 1.4e-66;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 433 GAGGAG 492
DB 455 GAGGAG 396
QY 493 GAGCAAGCCCTGTCAAAAGTCAAAAGCTGTGCTCCCAAGAGAGAGAGAGAGAG 552
DB 395 GAGCAAGCCCTGTCAAAAGTCAAAAGCTGTGCTCCCAAGAGAGAGAGAGAGAGAG 336
QY 553 AAGAAAAAG 612
DB 335 AAGAAAAAG 276
QY 613 CCTGTGAAAAAGGCAAGCCAGAGCCAGAGCCAGAGATTCAGAGAAATAGAGA 672
DB 275 CCTGTGAAAAAGGCAAGCCAGAGCCAGAGCCAGAGATTCAGAGAAATAGAGA 216
QY 673 GCCAG 732
DB 215 GCCAG 156
QY 733 GGTGAG 792
DB 155 GGTGAG 96
QY 793 CATGAG 852
DB 95 CATGAG 36
QY 853 CTCGGGGTCAACATTAAGTGTGCTGTCAGGAGAGAA 887
DB 35 CTCGGGGTCAACATTAAGTGTGCTGTCAGGAGAGAA 1

RESULT 3

REFERENCE
1 (bases 1 to 433)
Melton, D., Brown, J., Kenty, G., Permutt, A., Lee, C., Kaesinger, K.,

TITLE	Endocrine Pancreas Consortium
JOURNAL	Unpublished (2000)
COMMENT	Other_ESTS: 1e92b10.x1

Dr. Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
Endocrine Pancreas Consortium
Harvard University, Howard Hughes Medical Institute
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
MA 02138
Tel: 617-495-1812
Fax: 617-495-8557
Email: dmelton@biohp.harvard.edu
Library was constructed by Dr. Douglas Melton DNA sequencing by:
Washington University Genome Sequencing Center for Information on
obtaining a clone please contact: Juliana Brown
(brown@fas.harvard.edu) This sequence now available from the IMAGS
consortium. For clone orders contact: info@image.llnl.gov.
Location/Qualifiers

FEATURES	VALUES
Location/Qualifiers	intocmagc.intl.gov.
Source	1. .433

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="5674315"
/clone_lib="Mellon Normalized Human Islet 4 N4-HIS 1"
/sex="Both"
/issue_type="Islets of Langerhans"
/dev_stage="Adult"
/lab_host="DH10B"
/notes="Organ: Pancreas; Vector: pSPORT1; Site_1: Not 1; Site_2: Sal 1; Starting library constructed using SuperScript Plasmid Library Kit (Life Technologies). cDNA made by oligo-dT priming. Size-selected by column fractionation; average insert size 1.08 kb. Library was amplified once on solid support and plasmid DNA from library was prepared. The library DNA was normalized by method #4 from Bonaldo, Lennon, and Soares 1996 Genome Research 6:791-806; 0.5 microgram single-stranded library plasmid DNA was mixed with 5 micrograms PCR product representing library inserts and hybridized to an EcoT of 20. Single-stranded (unhybridized) plasmids were isolated by hydroxyapatite chromatography and used to make this library."

```

Query Match	46.9%	Score 433	DB 13	Length 433
Best Local Similarity	100.0%	Pred. No. 6.6e-63		
Matches 433	Conservative 0	Mismatches 0	Indels 0	Gaps 0
OY	391	GCATCAGACCTTAACCTGGGAGGAGGAGAGCAAGAAGGCGGAGGAGGAAGGAA	450	
Db	1	GCATTCAGACCTTAACCTGGGAGGAGGAGCAAGAGGGGAGGAGGAAGGAA	60	
OY	451	GAGCAAGATGATGAGGATGAGGATGCAGATATATCTCTGGAGAGCAAAACCCCTGTCAA	510	
Db	61	GAGCAAGATGATGAGGATGAGGATGCAGATATATCTCTGGAGAGCAAAACCCCTGTCAA	120	

Qy	511	CAATCAAAAGCGTGGGCCCGAAGCGAGGGAGCGTGGCTAAAGAAAAAAACCTGGA	570
Db	121	CAATCAAAAGCGTGGGCCCGAAGCGAGGGAGCGTGGCTAAAGAAAAAAACCTGGA	180
Qy	571	AAAGAGAGAGGAATATAGAGCCAGCGTTAGAGACAAGAGCCCTGTGTAAAAACGCCAA	630
Db	181	AAAGAGAGAGGAATATAGAGCCAGCGTTAGAGACAAGAGCCCTGTGTAAAAACGCCAA	240
Qy	631	GCCACAGCCAGAGCCCAAGAAGCGAGGATTCAGAAATAGAGAGCCAGCGTTGGGGGCA	690
Db	241	GCCACAGCCAGAGCCCAAGAAGCGAGGATTCAGAAATAGAGAGCCAGCGTTGGGGGCA	300
Qy	691	CGGTGCAAAATGGGCTTCCCTGAGCTGTGTCAGAGGAGGAGGTCCTGTGTCAAGCC	750
Db	301	CGGTGCAAAATGGGCTTCCCTGAGCTGTGTCAGAGGAGGAGGTCCTGTGTCAAGCC	360
Qy	751	CTCACCTGTGTGTAATGCAACAGAGGGGTGTTGCCGGGCAACATTGAGAGCCCTCAACC	810
Db	361	CTCACCTGTGTGTAATGCAACAGAGGGGTGTTGCCGGGCAACATTGAGAGCCCTCAACC	420
Qy	811	CCAACTCTCACCT 823	
Db	421	CCAACTCTCACCT 433	

RESULT 4	LOCUS	DEFINITION	CDNA CLONE IMAGE	EST	12-MAR-2002
BM054976/c	BM054976	451 bp	Normalised Human	1	1 Homo sapiens
ACCESSION	BM054976	1	1	1	1
VERSION	BM054976.1	1	1	1	1
KEYWORDS	EST.	1	1	1	1
SOURCE	human.	1	1	1	1

ORGANISM	<i>Homo sapiens</i> Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. (bases 1 to 451)
REFERENCE	Melton, D., Brown, J., Kenty, G., Permut, A., Lee, C., Kaestner, K., Lemisha, I., Sceaer, M., Brestelli, J., Girdwohl, G., Clifton, S., Hillier, L., Warr, M., Page, D., Wylie, T., Martin, J., Bliston, A., Schmitt, A., Theising, B., Riltner, E., Ronko, I., Bennett, J., Cardenas, M., Gibbons, M., McCann, R., Cole, R., Tsagariswili, R., Williams, T., Jackson, Y. and Bowers, Y.
TITLE	Endocrine Pancreas Consortium
JOURNAL	Unpublished (2000)
COMMENT	Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue

Harvard University, Howard Hughes Medical Institute
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
 MA 02138
 Tel: 617-495-1812
 Fax: 617-495-8557
 Email: dmelton@iobhp.harvard.edu
 Library was constructed by Dr. Douglas Melton DNA sequencing by:
 Washington University Genome Sequencing Center for information on
 obtaining a clone please contact: Juliana Brown
 (brown@fas.harvard.edu) This sequence now available from the IMAGE
 consortium, for clone orders contact: info@imgc.llnl.gov.
 Location/Qualifiers

```

source
1. 451
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="5674315"
/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"
/sex="Both"
/tissue_type="Islets of Langerhans"
/dev_stage="Adult"
/lab_host="DH10B"
/note="Organ: Pancreas; Vector: pSPORT1; Site:1; Not 1;
Site:2: Sal 1; Starting library constructed using
Superscript Plasmid library kit (Life Technologies). cDNA
made by oligo-dT priming. Size-selected by column

```


BO878692
 LOCUS BO878692 854 bp mRNA linear EST 16-AUG-2002
 DEFINITION AGENCOURT_8072876 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6087706
 5', mRNA sequence.
 ACCESSION BO878692
 VERSION BO878692.1 GI:22270700
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 854)
 NIH-MGC <http://mgc.ncl.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: c9apbs-remail.nih.gov
 Tissue Procurement: DCTD/DTP
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLCM2324 row: 1 column: 11
 High quality sequence stop: 560.
 Location/Qualifiers
 1..854
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="IMAGE:6087706"
 /clone_1lb="NIH_MGC_112"
 /tissue_type="melanotic melanoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GCCACGAC(G). Library constructed by Ling Hong in the
 Laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH_MGC library."
 BASE COUNT 181 a 247 c 253 g 151 t 22 others
 ORIGIN
 Query Match 44.8%; Score 414.4; DB 14; Length 854;
 Best Local Similarity 76.5%; Pred. No. 6.7e-60;
 Matches 595; Conservative 0; Mismatches 16; Indels 167; Gaps 1;
 Oy 1 CAGCCGCTCTCTGCGCCGACCATGATCTCACTACGCCAGTAGCAGGAGAAAG 60
 |||||||
 Db 207 CAGCCGCTCTCTGCGCCGACCATGATCTCACTACGCCAGTAGCAGGAGAAAG 266
 Oy 61 CGAGTGACGACCGCTCTGCGGCTGCGAGCTCAGAGAGAGCGGACTTGGACCTTC 120
 |||||||
 Db 267 CGAGTGACNNCGCTCTGCGGCTGCGAGGCTCAGTGCAGAGAGCGGACTTGGACCTTC 326
 Oy 121 AGACCCCACTGGAGGAGGAGAGAGAGCTGAGCTGTCTTCATACGATTGGCTGGG 180
 |||||||
 Db 327 AGACCCCACTGGAGGAGGAGAGAGCTGAGCTGTCTTCATACGATTGGCTGGG 386
 Oy 181 GAGAAAGCCAAAGAGAGATGCGCTGGAGATCCGCGCCGACGAAACCGAGAGAG 240
 |||||||
 Db 387 GAGAAAGCCAAAGAGAGATGCGNTGAGATCTCGCCCGACGAAACCGAGAGAG 446
 Oy 241 AAGAAATGACGACCGGCTACCACTTGCCTCACTCAGGCGCTCAGTCTCCCATGCTTC 300
 |||||||
 Db 447 AAGAAATGACGACCGGCTACCACTTGCCTCACTCAGGCGCTCAGTCTCCCATGCTTC 506
 Oy 301 ATGTAGAGAGAGAGCTTCTCCCGCAGTACTTCCACGCTCCGGCTGCTCAGAGACC 360
 |||||||
 Db 507 ATGTAGAGAGAGCTTCTCCCGCAGTACTTCCACGCTCCCGCAGTACTTCCAGCTCAGAGACC 566

Oy 361 GTGTTCTCACTGGCCAGAACGTTATGAAGCATCACTTAACCTTGGAGGAGAGAG 420
 |||||||
 Db 567 GTGTTCTCACTGGCCAGAACGTTAT----- 593
 Oy 421 GAAGAAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 480

 Db 594 ----- 593
 Oy 481 ATATCTGGAGAGCAAAAGCCCTGTCAAAAGTCAAAAGCTGTGCTCCCAAGACAG 540
 Db 594 ----- 593
 Oy 541 GCGAGCTGGCTAAGAAAAAGCTGAAAAAGAGAGAGAAATAGACGACGCTT 600
 |||||||
 Db 594 -----GAAAAAAAGCTGAAAAAGAGAAATAGACGACGCTT 639
 Oy 601 AGAGCAAGAGCCCTGTGAAAAAGGCAAGCCACAGCCAGCAAGAGCCAGGATTC 660
 |||||||
 Db 640 AGAGCAAGAGCCCTGTGAAAAAGGCTTNNCTCAGCCAGCAAGAGCCAGGATTC 699
 Oy 661 AAGAAATGAGAGCAAGCCCTTGGGGGCGACGGTGCAGAGTGGGCTTCCCTGGGCTGG 720
 |||||||
 Db 700 AAGAAATGAGAGCAAGCCCTTGGGGGCGACGGTGCAGAGTGGGCTTCCCTGGGCTGG 759
 Oy 721 CTGAGGCAAGAGGCTGCTCCCTGTCACGCCCCCTTCACCTGTGTGAATGCAACAGGG 778
 |||||||
 Db 760 CTGAGGCAAGAGGCTGCTCCCTGTCACGCCCCCTTCACCTGTGTGAATGCAACAGGG 817

RESULT 7
 BO684634
 LOCUS BO684634 915 bp mRNA linear EST 15-JUL-2002
 DEFINITION AGENCOURT_8032849 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6091012
 5', mRNA sequence.
 ACCESSION BO684634
 VERSION BO684634.1 GI:21797313
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 915)
 NIH-MGC <http://mgc.ncl.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: c9apbs-remail.nih.gov
 Tissue Procurement: DCTD/DTP
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLCM2333 row: f column: 05
 High quality sequence stop: 652.
 Location/Qualifiers
 1..915
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="IMAGE:6091012"
 /clone_1lb="NIH_MGC_112"
 /tissue_type="melanotic melanoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:
 EcoRI; cDNA made by oligo-dT priming. Directionally cloned
 into EcoRI/XhoI sites using the following 5' adaptor:
 GCCACGAC(G). Library constructed by Ling Hong in the
 Laboratory of Gerald M. Rubin (University of California,
 Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
 Superscript II RT (Life Technologies). Note: this is a
 NIH_MGC library."
 BASE COUNT 197 a 283 c 277 g 153 t 5 others

ORIGIN

Query Match 44.4%; Score 410.6; DB 14; Length 915;
 Best Local Similarity 77.0%; Pred. No. 2.8e-59;
 Matches 631; Conservative 0; Mismatches 17; Indels 172; Gaps 4;

QY 1 CAGCCCGCTTCTCTGCGGAGCCATGATCTCAGTAGACGCCGCTACGACGAGGAAAG 60
 DB 207 CAGCCCGCTTCTCTGCGGAGCCATGATCTCAGTAGACGCCGCTACGACGAGGAAAG 266
 QY 61 GCAATGACGACCGCTCTGGGGTGGAGCTAGTAGAGAGCGGACTTGGACCTTC 120
 DB 267 GCAATGACGACCGCTCTGGGGTGGAGCTAGTAGAGAGCGGACTTGGACCTTC 326
 QY 121 AGACCCGAGCTGAGGAGGAGAGAGAGCTGAGGCTGTTCCTTCAATGATTTGCTGGG 180
 DB 327 AGACCCGAGCTGAGGAGGAGAGAGAGCTGAGGCTGTTCCTTCAATGATTTGCTGGG 386
 QY 181 GAGAAAGCCAAAGAGAGATGATCGCTGAGATCTGCGCCCGACGAAACGAGAGAG 240
 DB 387 GAGAAAGCCAAAGAGAGATGATCGCTGAGATCTGCGCCCGACGAAACGAGAGAG 446
 QY 241 AAGAAAGATGAGCGGCTGACCATTCCTCAGCTCAGGCTCAGTCTCCCATGCTCC 300
 DB 447 AAGAAAGATGAGCGGCTGACCATTCCTCAGCTCAGGCTCAGTCTCCCATGCTCC 506
 QY 301 ATGTAGAGATGAGCTTTCTCCCGAGTTACTTTCCAGCTCGGGCTGGCTCAGGACC 360
 DB 507 ATGTAGAGATGAGCTTTCTCCCGAGTTACTTTCCAGCTCGGGCTGGCTCAGGACC 566
 QY 361 GTGTTCTCAGTGGCCAGAGAGCTTATGAGCATCAGACCTAACCTGGAGAGAGAG 420
 DB 567 GTGTTCTCAGTGGCCAGAGAGCTTATGAGCATCAGACCTAACCTGGAGAGAGAG 593
 QY 421 GAAGAAGAGGAGAGAGAGAGAGAGAGAGATGATGAGATGAGATGAGATGAGAT 480
 DB 594 ----- 593
 QY 481 ATATCTGAGAGAGCAAGCCCTCTCAAAAGTCAAAAGGCTGTGCTCCCGAGAGAG 540
 DB 594 ----- 593
 QY 541 GCGAGCGTGGCTAGAGAAAAAGCTGGAAGAAAGAGAGAAATTAAGAGCGCTT 600
 DB 594 -----GAAAAAAGCTGGAAGAAAGAGAGAAATTAAGAGCGCTT 639
 QY 601 AGAGACAGAGAGCGCTGTGAAAAAGGCCAAAGCCAGACGCCAGAGCAAGAGAGATTC 660
 DB 640 AGAGACAGAGAGCGCTGTGAAAAAGGCCAAAGCCAGACGCCAGAGCAAGAGATTC 698
 QY 661 AAGAAATGAGAGAGCGCTGTGAAAAAGGCCAAAGCCAGAGCGCTTCCCTGGGCTGTG 720
 DB 699 AAGAAATGAGAGAGCGCTGTGAAAAAGGCCAAAGCCAGAGCGCTTCCCTGGGCTGTG 758
 QY 721 CTGAGAGAGAGAGCGCTGTGAAAAAGGCCAAAGCCAGAGCGCTTCTGATATGCAACAGGGT 779
 DB 759 CTGAGAGAGAGAGCGCTGTGAAAAAGGCCAAAGCCAGAGCGCTTCTGATATGCAACAGGGT 818
 QY 780 --GTTGCGGAGGAGAGATG-AGAGCGCCCTCAGACCCCAACT 816
 DB 819 GTTTCGCGGAGGAGAGATG-AGAGCGCCCTCAGACCCCAACT 858

RESULT 8
 B0878312 872 bp mRNA linear EST 16-AUG-2002
 LOCUS B0878312
 DEFINITION AGENDCODET_8072779 NIH_MGC_1112 Homo sapiens CDNA clone IMAGE:6087652
 5', mRNA sequence.
 ACCESSION B0878312
 VERSION B0878312.1 GI:22270320
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS NIH-MGC
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: c9abbs-remail.nih.gov
 Tissue procurement: DCCD/DRP
 CDNA Library Preparation: Rubin Laboratory
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
 http://image.llnl.gov
 Plate: L1CM2324 row: j column: 05
 High quality sequence stop: 548.
 Location/Qualifiers

FEATURES
 source 1..872
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6087652"
 /clone_1ib="NIH_MGC_112"
 /tissue_type="melanotic melanoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pORF7; Site:1: XhoI; Site:2: EcoRI; CDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC Library."

BASE COUNT 192 a 265 c 259 g 145 t 11 others
 ORIGIN

Query Match 43.9%; Score 405.8; DB 14; Length 872;
 Best Local Similarity 76.1%; Pred. No. 1.8e-58;
 Matches 602; Conservative 0; Mismatches 20; Indels 169; Gaps 2;

QY 1 CAGCCCGCTTCTCTGCGGAGCCATGATCTCAGTAGACGCCGCTACGACGAGGAAAG 60
 DB 207 CAGCCCGCTTCTCTGCGGAGCCATGATCTCAGTAGACGCCGCTACGACGAGGAAAG 266
 QY 61 GCAATGACGACCGCTCTGGGGTGGAGCTAGTAGAGAGCGGACTTGGACCTTC 120
 DB 267 GCAATGACGACCGCTCTGGGGTGGAGCTAGTAGAGAGCGGACTTGGACCTTC 326
 QY 121 AGACCCGAGCTGAGGAGGAGAGAGAGCTGAGGCTGTTCCTTCAATGATTTGCTGGG 180
 DB 327 AGACCCGAGCTGAGGAGGAGAGAGAGCTGAGGCTGTTCCTTCAATGATTTGCTGGG 386
 QY 181 GAGAAAGCCAAAGAGAGATGATCGCTGAGATCTGCGCCCGACGAAACGAGAGAG 240
 DB 387 GAGAAAGCCAAAGAGAGATGATCGCTGAGATCTGCGCCCGACGAAACGAGAGAG 446
 QY 241 AAGAAAGATGAGCGGCTGACCATTCCTCAGCTCAGGCTCAGTCTCCCATGCTCC 300
 DB 447 AAGAAAGATGAGCGGCTGACCATTCCTCAGCTCAGGCTCAGTCTCCCATGCTCC 506
 QY 301 ATGTAGAGATGAGCTTTCTCCCGAGTTACTTTCCAGCTCGGGCTGGCTCAGGACC 360
 DB 507 ATGTAGAGATGAGCTTTCTCCCGAGTTACTTTCCAGCTCGGGCTGGCTCAGGACC 566
 QY 361 GTGTTCTCAGTGGCCAGAGAGCTTATGAGCATCAGACCTAACCTGGAGAGAGAG 420
 DB 567 GTGTTCTCAGTGGCCAGAGAGCTTATGAGCATCAGACCTAACCTGGAGAGAGAG 593
 QY 421 GAAGAAGAGGAGAGAGAGAGAGAGAGAGATGATGAGATGAGATGAGATGAGAT 480
 DB 594 ----- 593
 QY 481 ATATCTGAGAGAGCAAGCCCTCTCAAAAGTCAAAAGGCTGTGCTCCCGAGAGAG 540

Db 594 ----- 593

QY 541 GCGAGCTGGCTAGAAAAAGCTGGAAAAAGAAAGAGAGAAATAGAGCCAGCTT 600

Db 594 ----- GAAAAAAGCTGGAAAAAGAAAGAGAAATAGAGCCAGCTT 639

QY 601 AGACACAGAGCCCTGTGAAAAAGGCCAAAGCCACAGCCAGAGCCAAAGAGCCAGATTTC 660

Db 640 AGAGACAAAGAGCCCTGTGAAAAAGGCCAAAGCCACAGCCAGAGCCAAAGAGCCAGATTTC 699

QY 661 AAGAAATGAGAGCCAGAGCCCTTGGGGGACAGCGTGCAGAAAGTGGCCCTTGGGCTGTG 720

Db 700 AAGAAATGAGAGCCAGAGCCCTTGGGGGACAGCGTGCAGAAAGTGGCCCTTGGGCTGTG 759

QY 721 CTCGAGCAGAGCCCTGTGCAAGCCCTTCCAGCTGTCTGAATGCAACAGGG 778

Db 760 CTCGAGCAGAGCCCTGTGCAAGCCCTTCCAGCTGTCTGAATGCAACAGGG 819

QY 779 TGTTCGCGGGG 789

Db 820 GGTTCGCGGG 830

RESULT 9
BF594409/c 405 bp mRNA linear EST 12-DEC-2000

LOCUS 7105a05.x1 NCI_CGAP.Col6 Homo sapiens cDNA clone IMAGE:3324560.3'

DEFINITION similar to contans Alu repetitive element; contains element MER28

ACCESSION BF594409

VERSION BF594409.1 GI:11686733

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 405)

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index

JOURNAL Unpublished (1997)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
Tissue Procurement: Ian Kirsch, M.D., Michael R. Emmert-Buck, M.D., Ph.D.

cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL, send email to: info@image.lnl.gov
Seq primer: -400p from Glibco
High quality sequence stop: 404.

FEATURES
source
1..405
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="IMAGE:3324560"
/clone_id="NCI_CGAP.Col6"
/tissue_type="Colon tumor, RER"
/lab_host="DH10B"

/note="Organ: colon; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA from the normalized library NCI CGAP.Col6 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clones 1057416-1061255, and 1144584-1145351). Subtraction by Bento Soares and M. Fatima Bonaldo. "

BASE COUNT 54 a 120 c 107 g 124 t

ORIGIN

Query Match 43.8%; Score 405; DB 12; Length 405;
Best Local Similarity 100.0%; Pred. No. 3.2e-58;
Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 488 TGGAGAGCAAAAGCCCTGTCTAAACAGTCAAAAGCTGTGCCCCAGAGAGCGGAGCG 547

Db 405 TGGAGAGCAAAAGCCCTGTCTAAACAGTCAAAAGCTGTGCCCCAGAGAGCGGAGCG 346

QY 548 TGGCTAAGAAAAAGCTGGAAGAAAGAGAGAAATTAAGAGCCAGCTTAAAGACA 607

Db 345 TGGCTAAGAAAAAGCTGGAAGAAAGAGAGAAATTAAGAGCCAGCTTAAAGACA 286

QY 608 AGAGCCCTGTGAAAAAGGCCAAAGCCACAGCCAGAGCCAAAGAGCCAGATTCAAGAAAT 667

Db 285 AGAGCCCTGTGAAAAAGGCCAAAGCCACAGCCAGAGCCAAAGAGCCAGATTCAAGAAAT 226

QY 668 GAGAGCCAGAGCCCTTGGGGGGGACAGGTGCAAGTGGCCCTTCCCTGCTGTGCTGACAG 727

Db 225 GAGAGCCAGAGCCCTTGGGGGGGACAGGTGCAAGTGGCCCTTCCCTGCTGTGCTGACAG 166

QY 728 CACAGGGTGGCCCTGTGCAAGCCCTTCCAGCTGTCTGATGCAACAGAGGGGTGTCGAGG 787

Db 165 CACAGGGTGGCCCTGTGCAAGCCCTTCCAGCTGTCTGATGCAACAGAGGGGTGTCGAGG 106

QY 788 GCGAATGAGAGCCCTTCAACCCCACTCTGCTGAGAGGGGCCAGTGAAGAGC 847

Db 105 GCGAATGAGAGCCCTTCAACCCCACTCTGCTGAGAGGGGCCAGTGAAGAGC 46

QY 848 CCCACCTGGGGTCACAATTAAGTGGCGGTGACGAAAAA 892

Db 45 CCCACCTGGGGTCACAATTAAGTGGCGGTGACGAAAAA 1

RESULT 10
AW731946/c 399 bp mRNA linear EST 21-APR-2000

LOCUS ba03e06.x1 NIH_MGC_7 Homo sapiens cDNA clone IMAGE:2823298.3', mRNA

DEFINITION sequence.

ACCESSION AW731946

VERSION AW731946.1 GI:7632268

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 399)

AUTHORS NIH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
cDNA Library Preparation: Ling Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLU)
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: image.lnl.gov/image/ncmi/lresources.shtml
Seq primer: -400p from Glibco
High quality sequence stop: 381.

FEATURES
source
1..399
location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_image="IMAGE:2823298"
/clone_id="NIH_MGC_7"
/tissue_type="small cell carcinoma"
/cell_line="MGC3"
/lab_host="DH10B (phage-resistant)"

/note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GCCACGAG(C). Size-selected >500bp for average

Insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."

BASE COUNT 53 a 116 c 108 g 122 t

ORIGIN

Query Match 43.0%; Score 397.4; DB 10; Length 399;
Best Local Similarity 99.7%; Pred. No. 5.9e-57;
Matches 398; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

493 GAGCAAGCCCTGTCACAAAGTCACAAAGGCTGGTCCCGACAGAGCGGCGCTGGCT 552
|||||
Db 399 GAGCAAGCCCTGTCACAAAGTCACAAAGGCTGGTCCCGACAGAGCGGCGCTGGCT 340
553 AAGCAAAAAAGCTGCAAAAAAGAGAAATTAAGACCGCTTGAAGCAAGAGC 612
|||||
Db 339 AAGCAAAAAAGCTGCAAAAAAGAGAAATTAAGACCGCTTGAAGCAAGAGC 280
613 CCTGTGAAAAAGGCAAGGCGACAGAGCCAGAGGCGGATTCAGAAATGAGGA 672
Db 279 CCTGTGAAAAAGGCAAGGCGACAGAGCCAGAGGCGGATTCAGAAATGAGGA 220
673 GCGACGCTTGGGGGCGACGCTGCAAAAGTGGCTTCCCTGGGCTGTGCTGCAGAGC 732
Db 219 GCGACGCTTGGGGGCGACGCTGCAAAAGTGGCTTCCCTGGGCTGTGCTGCAGAGC 160
733 GGTGCCCCCTGTCAGCCCTCCACCTGTGCTGAATGCAACAGAGGCTGTGGCGGGCAA 792
Db 159 GGTGCCCCCTGTCAGCCCTCCACCTGTGCTGAATGCAACAGAGGCTGTGGCGGGCAA 100
793 CATGAGAGCCCTCCACCCCACTCTCCACTTTAGAGAGGCGCCCACTGAAGAGCCAC 852
Db 99 CATGAGAGCCCTCCACCCCACTCTCCACTTTAGAGAGGCGCCCACTGAAGAGCCAC 40
853 CTCGGGCTCAATTAAGTTCCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 891
|||||
Db 39 CTCGGGCTCAATTAAGTTCCTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAG 1

RESULT 11
B0684424 935 bp mRNA linear EST 15-JUL-2002
LOCUS AGENCOURT_8209062 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6260492
DEFINITION 5', mRNA sequence.

ACCESSION B0684424
VERSION B0684424.1 GI:21797103
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 935)
NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1CM2419 row: k column: 21
High quality sequence stop: 635.

FEATURES
SOURCE

1. 935
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:6260492"
/clone_lib="NIH_MGC_112"

/tissue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pOTB1; Site:1: XhoI; Site:2:
EcoRI; cDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GCCACGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."

BASE COUNT 186 a 304 c 269 g 175 t 1 others
ORIGIN

Query Match 42.2%; Score 389.6; DB 14; Length 935;
Best Local Similarity 98.7%; Pred. No. 8.9e-56;
Matches 392; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

1 CAGCCCGCTTCTCTGCGGAGCCATGATCTCAGTACGCGCAGTACGAGGAGAAAG 60
|||||
Db 207 CAGCCCGCTTCTCTGCGGAGCCATGATCTCAGTACGCGCAGTACGAGGAGAAAG 266
61 GCGATGACGACCGTCTGCTGGGCTTGGAGCTCAGTCAGAGAGAGGCGGACTTGGACCTTC 120
|||||
Db 267 GCGATGACGACCGTCTGCTGGGCTTGGAGCTCAGTCAGAGAGGCGGACTTGGACCTTC 326
121 AGACCCGCTTGGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 180
Db 327 AGACCCGCTTGGAGGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 386
181 GAGAAAGCCAAAG 240
Db 387 GAGAAAGCCAAAG 446
241 AAGAAATGAGAGGCGGCTGACCATTCCTCAGTCAGAGAGAGAGAGAGAGAGAGAGAG 300
Db 447 AAGAAATGAGAGGCGGCTGACCATTCCTCAGTCAGAGAGAGAGAGAGAGAGAGAGAG 506
301 ATGTAGAGAGTGCAGGCTTCTCCCGAGTACTTTCAGCTCCGAGGCTGCTCAGAGAGCC 360
Db 507 ATGTAGAGAGTGCAGGCTTCTCCCGAGTACTTTCAGCTCCGAGGCTGCTCAGAGAGCC 566
361 GTGTCTCTCAGTGGCCAG 397
|||||
Db 567 GTGTCTCTCAGTGGCCAG 603

RESULT 12
B0682257 966 bp mRNA linear EST 15-JUL-2002
LOCUS AGENCOURT_8194971 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:6259047
DEFINITION 5', mRNA sequence.

ACCESSION B0682257
VERSION B0682257.1 GI:21794936
KEYWORDS EST.

SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 966)
NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: c9apbs-remail.nih.gov
Tissue Procurement: DCTD/DTF
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: L1CM2415 row: o column: 16
High quality sequence stop: 518.

FEATURES
sourceLocation/Qualifiers
1. 966

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5259047"
/clone_lib="NIH_MGC_112"
/issue_type="melanotic melanoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pORF7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCCAGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH_MGC library."

BASE COUNT 205 a 300 c 265 g 159 t 7 others
ORIGIN

Query Match 41.0%; Score 378.4; DB 14; Length 966;
Best Local Similarity 74.0%; Pred. No. 6.5e-54;
Matches 668; Conservative 0; Mismatches 56; Indels 179; Gaps 8;

QY 1 CACCCGCTTCTCTGCGCCGACCATGATCTCACTAGCCGACGATGACGAGGAAAG 60
DB 207 CACCCGCTTCTCTGCGCCGACCATGATCTCACTAGCCGACGATGACGAGGAAAG 266
QY 61 GCAGTACGACGCTGCTGGGGCTGCGAGCTCACTAGGAGGAGCGGACTTGGACCTTC 120
DB 267 GCAGTACGACGCTGCTGGGGCTGCGAGCTCACTAGGAGGAGCGGACTTGGACCTTC 326
QY 121 AGACCCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 180
DB 327 AGACCCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 386
QY 181 GAGAAAGCCAAAGAGAGAGATGC-ATCGGCTGGAATCTCTGCCCCAGCAACAGAGAGA 239
DB 387 GAGAAAGCCAAAGAGAGATGC-ATCGGCTGGAATCTCTGCCCCAGCAACAGAGAGA 446
QY 240 CAAGAAGATGACGCGCTGACCATGCTCACTCAAGCCGCTCACTGCTCCCATAGTCTC 259
DB 447 CAAGAAGATGACGCGCTGACCATGCTCACTCAAGCCGCTCACTGCTCCCATAGTCTC 506
QY 300 CATGG--TAGGAGTCACTCTTCTCCCACTTCACTTCACTTCACTTCACTTCACT 356
DB 507 CATGGTGGGAGTGCACCTTCTCTCCCACTTCACTTCACTTCACTTCACTTCACT 566
QY 357 -ACCCGCTGCTCTAGTGGCGAGAGGCTTGAAGCATGACGACTTAACCTGGGAGGAG 415
DB 567 ACCCGTGGTCTCTAGTGGCGAGAGGCTTGAAGCATGACGACTTAACCTGGGAGGAG 598
QY 416 AGGAGGAAGAAAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 475
DB 599 ----- 598
QY 476 CAGATATATCTCTGAGAGAGCAAAAGCCCTGTCAAAACAGTCAAAAGAGCTGGTCCCA 535
DB 599 ----- 598
QY 536 AGCAGCGCAGCGTGGCTTAAGAAAAAGCTGAAAAAGCAAAAGCCACAGCGACAGCC 595
DB 599 -----GAAAAAAGCTGAAAAAGCAAAAGAGGAGGAGGAGGAGGAGGAGGAG 639
QY 596 GCGTTAGAGCAAGAGCCCTGTGAAAAAGCCAAAGCCACAGCGACAGCCAGCCAGAGCC 655
DB 640 GCGTTAGAGCAAGAGCCCTGTGAAAAAGCCAAAGCCACAGCGACAGCCAGAGCCAG 699
QY 656 GATTCAAGAAATGAGAGCCAC--GCCTTGGGGGAGGAGGAGGAGGAGGAGGAGGAGG 712
DB 700 GATTCAAGAAATGAGAGCCCGGCTTGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAG 759
QY 713 GCGCTGTCTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 772

DB 760 GCGCTGTCTGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 819
QY 773 CAGGCGGTG--TTGCGGGGGGACATGAGAGCCCTCACCCGACCCGACCTTCATGACA 830
DB 820 CAGGCGGTGTTGGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 879
QY 831 -GGCCCCGATGAGAGCCGACCTGCGGGGCTGACAAATTAAGTTCGTCAGGAGAA 888
DB 880 AGGCGCCGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 939
QY 889 AAA 891
DB 940 AAA 942

RESULT 13
AM002370/c 361 bp mRNA linear EST 09-MAR-2000
LOCUS
DEFINITION
wu61a07.x1 NCI-CGAP-GC6 Homo sapiens cDNA clone IMAGE:2524500 3',
mRNA sequence.
ACCESSION
AM002370
VERSION
AM002370.1 GI:5849286
KEYWORDS
EST.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 361)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
NATIONAL Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
JOURNAL
COMMENT
CONTACT: Robert Strausberg, Ph.D.
Email: cgapdb-remail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
Bonaldo, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bdrrp/image/image.html
Insert length: 431 Std Error: 0.00
Seq primer: -400P from Gibco.

FEATURES
sourceLocation/Qualifiers
1. 361

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:2524500"
/clone_lib="NCI-CGAP-GC6"
/issue_type="pooled germ cell tumors"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker. Site_1: Not I; Site_2: Eco RI; Plasmid DNA
from the normalized library NCI-CGAP-GC4 was prepared, and
ss circles were made in vitro. Following HAP purification,
this DNA was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from a pool
of 5,000 clones made from the same library (clones
1257096-1258631, 1469064-1470983, and 1475592-1476743).
Subtraction by Bento Soares and M. Fatima Bonaldo."

BASE COUNT 50 a 107 c 93 g 111 t
ORIGIN

Query Match 38.7%; Score 357.8; DB 10; Length 361;
Best Local Similarity 99.4%; Pred. No. 2.5e-50;
Matches 359; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 533 AGAGCAGGCGAGGCTGCTTAAGAAAAAGCTGAAAAAGAGAGGAGGAGGAGGAGG 592
DB 361 AGAGCAGGCGAGGCTGCTTAAGAAAAAGCTGAAAAAGAGAGGAGGAGGAGGAGG 302

QY 593 CCAGCGTTAGAGACAGAGCCCTGTGAAAAAGGCCAACGACAGCCAGCAAGAAGC 652
 301 CCAGCGTTAGAGACAGAGCCCTGTGAAAAAGGCCAACGACAGCCAGCAAGAAGC 242
 QY 653 CAGGATTCAAGAAATAGAGAGCCAGCCCTTGGGGGGGCAAGGCAAGTGGGCTCCCT 712
 241 CAGGATTCAAGAAATAGAGAGCCAGCCCTTGGGGGGGCAAGGCAAGTGGGCTCCCT 182
 QY 713 GGGCTGTGTCAGAGCAGAGGGGCGCCCTGTCTCAGCGCCCTCCACTGTCTGAATGCAA 772
 181 GGGCTGTGTCAGAGCAGAGGGGCGCCCTGTCTCAGCGCCCTCCACTGTCTGAATGCAA 122
 QY 773 CAGGGGTGTGCGGGGGCAACATGAGAGCCCTCACCCCAACTCTCCACTTTGAGAGG 832
 121 CAGGGGTGTGCGGGGGCAACATGAGAGCCCTCACCCCAACTCTCCACTTTGAGAGG 62
 QY 833 CCCCCAGTGAAGAGCCCGACCTCGGGGTCAATTAAGTTGGCTGTGTCAGAAAAA 892
 61 CCCCCAGTGAAGAGCCCGACCTCGGGGTCAATTAAGTTGGCTGTGTCAGAAAAA 2
 Db 893 A 893
 Db 1 A 1
 RESULT 14
 BF592761 361 bp mRNA linear EST 12-DEC-2000
 LOCUS 719404.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:3442494 3'
 DEFINITION mRNA sequence.
 ACCESSION BF592761
 VERSION BF592761.1 GI:11685085
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 361)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 JOURNAL
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
 R. Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
 Bonaldo, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL, send email to:
 info@image.llnl.gov
 Seq primer: -400p from Gibco.
 Location/Qualifiers
 1..361
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3442494"
 /clone_lib="NCI_CGAP_GC6"
 /tissue_type="pooled germ cell tumors"
 /lab_host="DH10B"
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA
 from the normalized library NCI-CGAP_GC6 was prepared, and
 ss circles were made in vitro. Following HAP purification,
 this DNA was used as tracer in a subtractive hybridization
 reaction. The driver was PCR-amplified cDNAs from a pool
 of 5,000 clones made from the same library (clones IDs
 1257096-1258631, 1469064-1470983, and 1475592-1476743).
 Subtraction by Bento Soares and M. Fatima Bonaldo."
 BASE COUNT 50 a 106 c 94 g 111 t
 ORIGIN

Query Match 38.7%; Score 357.8; DB 12; Length 361;
 Best Local Similarity 99.4%; Pred. No. 2,5e-50;
 Matches 359; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
 QY 533 AGAAGCAGGCGAGCGCTGCTAAGAAAAAAGCTGGAAGAAAGAGAGGAATTAAGAG 592
 361 AGAAGCAGGCGAGCGCTGCTAAGAAAAAAGCTGGAAGAAAGAGAGGAATTAAGAG 302
 Db 593 CCAGCGTTAGAGACAGAGCCCTGTGAAAAAGGCCAACGACAGCCAGCAAGAAGC 652
 301 CCAGCGTTAGAGACAGAGCCCTGTGAAAAAGGCCAACGACAGCCAGCAAGAAGC 242
 QY 653 CAGGATTCAAGAAATAGAGAGCCAGCCCTTGGGGGGGCAAGGCAAGTGGGCTCCCT 712
 241 CAGGATTCAAGAAATAGAGAGCCAGCCCTTGGGGGGGCAAGGCAAGTGGGCTCCCT 182
 QY 713 GGGCTGTGTCAGAGCAGAGGGGCGCCCTGTCTCAGCGCCCTCCACTGTCTGAATGCAA 772
 181 GGGCTGTGTCAGAGCAGAGGGGCGCCCTGTCTCAGCGCCCTCCACTGTCTGAATGCAA 122
 QY 773 CAGGGGTGTGCGGGGGCAACATGAGAGCCCTCACCCCAACTCTCCACTTTGAGAGG 832
 121 CAGGGGTGTGCGGGGGCAACATGAGAGCCCTCACCCCAACTCTCCACTTTGAGAGG 62
 QY 833 CCCCCAGTGAAGAGCCCGACCTCGGGGTCAATTAAGTTGGCTGTGTCAGAAAAA 892
 61 CCCCCAGTGAAGAGCCCGACCTCGGGGTCAATTAAGTTGGCTGTGTCAGAAAAA 2
 Db 893 A 893
 Db 1 A 1
 RESULT 15
 BF057162/c 361 bp mRNA linear EST 16-OCT-2000
 LOCUS 7K16g07.x1 NCI_CGAP_GC6 Homo sapiens cDNA clone IMAGE:3444661 3'
 DEFINITION mRNA sequence.
 ACCESSION BF057162
 VERSION BF057162.1 GI:10811058
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 361)
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
 Tumor Gene Index
 Unpublished (1997)
 JOURNAL
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaps-remail.nih.gov
 Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
 R. Emmert-Buck, M.D., Ph.D.
 CDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
 Bonaldo, Ph.D.
 CDNA Library Arrayed by: Greg Lennon, Ph.D.
 DNA Sequencing by: Washington University Genome Sequencing Center
 Clone distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL, send email to:
 info@image.llnl.gov
 Seq primer: -400p from Gibco.
 Location/Qualifiers
 1..361
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:3444661"
 /clone_lib="NCI_CGAP_GC6"
 /tissue_type="pooled germ cell tumors"
 /lab_host="DH10B"
 /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; Plasmid DNA

from the normalized library NCI-GAP_G4 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (clonids 1257096-1258631, 1469064-1470983, and 1475592-1476743). Subtraction by Bento Soares and M. Fatima Bonaldo.

BASE COUNT 50 a 107 c 93 g 111 t
ORIGIN

Query Match 38.7%; Score 357.8; DB 12; Length 361;
Best Local Similarity 99.4%; Pred. No. 2.5e-50;
Matches 359; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 533 AGAAGCAGCGCAGCGTGGCTAAGAAAAAGCTGAAAAAGAGAAAGAAATTAAG 592
   |||||||
Db 361 AGAAGCAGCGCAGCGTGGCTAAGAAAAAGCTGAAAAAGAGAAAGAAATTAAG 302
   |||||||
QY 593 CCAGCGTTAGAGACAAGAGCCCTGTGAAAAAGGCCAAGCCAGAGCCAAAGAGC 652
   |||||||
Db 301 CCAGCGTTAGAGACAAGAGCCCTGTGAAAAAGGCCAAGCCAGAGCCAAAGAGC 242
   |||||||
QY 653 CAGGATTCAAGAAATGAGAGCCAGCCCTTTGGGGGCGACAGTGCAAAAGTGGCCTTCCCT 712
   |||||||
Db 241 CAGGATTCAAGAAATGAGAGCCAGCCCTTTGGGGGCGACAGTGCAAAAGTGGCCTTCCCT 182
   |||||||
QY 713 GGGCTGTGCTGACAGGCACAGGGTGCCTGTCCAGCCCTCCACCTGTGCTGAATGCA 772
   |||||||
Db 181 GGGCTGTGCTGACAGGCACAGGGTGCCTGTCCAGCCCTCCACCTGTGCTGAATGCA 122
   |||||||
QY 773 CAGGGGTGTTGGGGGGCAACATGAGAGCCCTCACCCCAACTCTCCACTTTCAAGAG 832
   |||||||
Db 121 CAGGGGTGTTGGGGGGCAACATGAGAGCCCTCACCCCAACTCTCCACTTTCAAGAG 62
   |||||||
QY 833 CCCCAGTGAAGAGCCCCACCTCGGGGTCAACAATAAGTTGCTGTCAGGAAAAA 892
   |||||||
Db 61 CCCCAGTGAAGAGCCCCACCTCGGGGTCAACAATAAGTTGCTGTCAGGAAAAA 2
   |||||||
QY 893 A 893
   |
Db 1 A 1

```

Search completed: January 24, 2003, 18:43:01
Job time : 2248 secs

